

STIC-Biotech/ChemLib

CRFE

155312 6/7 06h +02p

From: Epps-Ford, Janet
Sent: Thursday, June 02, 2005 5:21 PM
To: STIC-Biotech/ChemLib
Subject: 09/155,676

76
570

6/8 8 (40 hrs)

Please search

SEQ ID NOs: 1, 3, 4, and 6 in all pending and commercial nucleic acid sequence databases. Reverse translation search in all amino acid pending and commercial databases.

SEQ ID NO: 2, 5, 7 in all pending and commercial amino acid databases.

Thanks,

Janet L. Epps-Ford, Ph.D.

Art Unit 1635

Mailbox: Remsen 2C18

Office: Remsen 2C05

Phone: 571-272-0757

Fax: 571-273-0757

1 na 1906
3 na 2631
4 na 1253
6 na 4596
2 aa 604
5 aa 417
7 aa 947

RECEIVED
JUN - 3 2005
(STIC)

me

STAFF USE ONLY

Searcher: Port
Searcher Phone: 2-
Date Searcher Picked up: 6/7
Date Completed: 6/13
Searcher Prep/Rev. Time: 20
Online Time: 12

Type of Search

NA#: 4 AA#: 3
Interference: ✓ SPDI: ✓
S/L: ✓ Oligomer: ✓
Encode/Transl: ✓
Structure#: ✓ Text: ✓
Inventor: ✓ Litigation: ✓

Vendors and cost where applicable

STN: ✓
DIALOG: ✓
QUESTEL/ORBIS: ✓
LEXIS/NEXIS: ✓
SEQUENCE SYSTEM: ✓
WWW/Internet: ✓
Other(Specify): ✓

NO PAGE BLANK (USPTO)



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number 155312

TO: Janet Epps-Ford
Location: rem/2C05/2C18
Art Unit: 1635
Tuesday, June 14, 2005

Case Serial Number: 09/155676

From: Toby Port

Location: Biotech-Chem Library
REM1-A59

Phone: 272-2523

toby.port@uspto.gov

Search Notes

Dear Examiner Epps-Ford,

Here are the results of your search.

Please feel free to contact me if you have any questions.

Toby Port

THIS PAGE BLANK (USPTO)



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact ***the searcher or contact:***

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop-off or send completed forms to STIC Biotech-Chem Library, Remsen Bldg.



"HIS PAGE BLANK (USPTO)"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 22:58:14 ; Search time 647.812 Seconds
(without alignments)
17417.120 Million cell updates/sec

Title: US-09-155-676B-1
Perfect score: 1906
Sequence: 1 cattgggtcacgcggtggcg.....tcattccttgtaaaaaaaa 1906

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1899	99.6	1906	2 AAV03323	AAV03323 5' end of
2	1742	91.4	1782	4 AAD04333	Aad04333 Human TNF
3	1555	81.6	2873	4 AAD04335	Aad04335 Human TNF
4	1555	81.6	3139	4 AAD04334	Aad04334 Human TNF
5	1138.2	59.7	2248	4 AAH15665	Aah15665 Human CDN
6	792.8	41.6	6045	4 AAH18235	Aah18235 Human CDN
7	672.2	35.3	730	4 AAH04973	Aah04973 Human CDN
8	352.2	18.5	33147	4 AAK67282	Aak67282 Human imm
9	274.6	14.4	454	5 AAS64410	Aas64410 DNA encod
10	263.8	13.8	280	6 ABS72233	Abs72233 Human gen
11	215.8	11.3	542	12 ACH75347	Ach75347 Human gen
12	212.8	11.2	284	12 ACH89047	Ach89047 Human gen
13	195.6	10.3	463	5 AAS92200	Aas92200 DNA encod
14	135.4	7.1	474	5 AAS73549	Aas73549 DNA encod
15	133.4	7.0	190	3 AAC11927	Aac11927 Human sec
16	128	6.7	500	12 ACH75141	Ach75141 Human gen
17	124	6.5	124	12 ACH88841	Ach88841 Human gen
18	119	6.2	911	5 AAS64411	Aas64411 DNA encod
19	83.6	4.4	100	3 AAC12137	Aac12137 Human sec
20	66.8	3.5	1796	4 ABL09365	Ab109365 Drosophil

C 21	66.8	3.5	3796	4 ABL09364	Ab109364 Drosophil
C 22	63.8	3.3	3931	4 ABA09172	Aba09172 Human BG3
C 23	61.6	3.2	1253	2 AAV03325	AAV03325 Sequence
C 24	61	3.2	2000	8 ADA71938	Ada71938 Rice gene
C 25	60	3.1	214520	10 ADL13471	Adl13471 Osteoarth
C 26	59.6	3.1	263	6 ABL86718	Ab186718 Human ova
C 27	59.4	3.1	5684	3 AAC59953	Aac59953 Human sec
C 28	59.4	3.1	5684	8 ABZ73290	Abz73290 Secreted
C 29	59.4	3.1	5684	10 ABZ66903	Abz66903 Human sec
C 30	59	3.1	3927	10 ADF81729	Adf81729 Leukaemia
C 31	58.4	3.1	595	13 ADQ53383	Adq53383 Novel can
C 32	56.6	3.0	729	10 ADE07225	Ade07225 Novel cod
C 33	54.6	2.9	371	9 ACH37072	Ach37072 Human end
C 34	54.6	2.9	446	6 ABN95020	Abn95020 Gene #151
C 35	54.6	2.9	769	4 AA195747	Aa195747 Human neu
C 36	54.4	2.9	813	5 AAS85362	Aas85362 DNA encod
C 37	54	2.8	464	9 ACH27965	Ach27965 Human adu
C 38	53.2	2.8	476	7 ADS31085	Ads31085 Human gen
C 39	53	2.8	143	4 AAD04332	Aad04332 Human IRE
C 40	52.2	2.7	941	5 ABA21208	Aba21208 Human ner
C 41	52.2	2.7	941	5 ABA21210	Aba21210 Human ner
C 42	52.2	2.7	961	12 ADL12492	Adl12492 Human ste
C 43	52.2	2.7	3916	13 ADR07577	Adr07577 Full leng
C 44	51.8	2.7	617	5 ABA14117	Ab14117 Human ner
C 45	50.4	2.6	7093	6 ABZ35416	Abz35416 Human gen

ALIGNMENTS

RESULT 1
AAV03323
ID AAV03323 standard; cDNA; 1906 BP.
AC AAV03323;
XX
DT 15-APR-1998 (first entry)
DE 5' end of clone 9, which encodes a TRAF2 binding protein.
KW Human tumour necrosis factor receptor-associated factor 2; TRAF2;
KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;
KW intracellular signalling activity; acute hepatitis;
KW autoimmune-induced cell death; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 94..1906
FT /tag= a
FT /codon= 127-129, aa:Ser
FT /note= "no ATG start or STOP codon given"
XX
PN WO737016-A1.
XX
PD 09-OCT-1997.
XX
PF 01-APR-1997; 97WO-IL000117.
XX
PR 02-APR-1996; 96IL-00117800.
PR 26-AUG-1996; 96IL-00119133.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;
XX
DR WPI; 1997-503101/46.
DR P-PSDB; AAW42400.
XX
PT DNA encoding tumour necrosis factor receptor-associated factor binding
PT molecule - used for modulation or mediation in cells of the activity of
PT NF-KB.
XX

Qy 1801 TAGTCTCATTTGAGCTCCTCGATCCAGTCTTTCTGAAGCTGTGTTCTCTCGACTTTT 1860
Db 1801 TAGTCTCATTTGAGCTCCTCGATCCAGTCTTTCTGAAGCTGTGTTCTCTCGACTTTT 1860
Qy 1861 CATGTATGTGAGCAATAATTCCTTTCATTCTTCAATTCCTTGAATAAAAAA 1906
Db 1861 CATGTATGTGAGCAATAATTCCTTTCATTCTTCAATTCCTTGAATAAAAAA 1906

RESULT 2
AAD04333
ID AAD04333 standard; cDNA; 1782 BP.
XX
AC AAD04333;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human TNF receptor-associated factor (TRAF2) binding protein, IREN cDNA.
XX
KW Human; Tumour Necrosis Factor; TNF; TNF Receptor Associated Factor;
KW TRAF2; TRAF2 binding protein; IREN; IkappaB Regulator; immunosuppressive;
KW nuclear factor-kappaB; NF-kappaB; cytoskeletal; tumour; AIDS;
KW acquired immune deficiency syndrome; rheumatic disease; apoptosis;
KW autoimmune disease; septic shock; graft-vs-host reaction; inflammation;
KW anorexia; anti-HIV; therapy; ss.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..1626
FT FT /*tag= a
FT FT /product= "Human TRAF2-binding protein, IREN"
XX
XX WO200116314-A1.
XX
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-IL000517.
XX
XX
PR 02-SEP-1999; 99IL-00131719.
XX
XX (YEDA) YEDA RES & DEV CO LTD.
XX
XX Wallach D, Malinin N, Sinha I, Leu S;
XX
DR WPI; 2001-281387/29.
DR P-PSDB; AAE00683.
XX
XX
PT New DNA sequence encoding Tumor Necrosis Factor receptor associated
PT factor (TRAF) binding proteins (IREN) for treatment or prevention of
PT pathological conditions associated with NF-kappaB induction.
XX
XX
PS Claim 1; Fig 3B; 118pp; English.

The present cDNA sequence encodes human tumour necrosis factor (TNF) receptor-associated factor (TRAF2) binding protein, IREN. A fragment of this IREN molecule is capable of binding to TRAF2 protein at position 225-501. The invention relates to human tumour necrosis factor (TNF) receptor-associated factor (TRAF2) binding protein designated as IREN (IkappaB Regulator), its isoforms IREN-10B, IREN-E and their corresponding cDNA molecules. IREN is useful for modulating/mediating the activity of transcription factor NF (Nuclear Factor)-kappaB or any other intracellular signalling activity mediated by TRAF2. IREN is useful in the prevention and treatment of a pathological condition associated with NF-kappaB induction (abnormal) e.g. AIDS (acquired immune deficiency syndrome), autoimmune diseases, tumours, rheumatic diseases, anorexia, septic shock and graft-vs-host reactions. IREN also plays an important role in the control of inflammation and other non-apoptotic effects of TNF as well as in the control of apoptosis. The invention also relates to method for screening, identifying and producing a molecule capable of modulating activities mediated by IREN. IREN antibodies are useful for the purification of new proteins from different sources, including cell

CC extracts or transformed cell lines, in addition IREN can be used in
CC diagnostic purposes for identifying disorders related to abnormal
CC functioning of cellular effects mediated directly by TRAF proteins
XX
SQ Sequence 1782 BP; 456 A; 458 C; 504 G; 364 T; 0 U; 0 Other;
Query Match 91.4%; Score 1742; DB 4; Length 1782;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1775; Conservative 0; Mismatches 4; Indels 3; Gaps 3;
Qy 118 ATGAGCGGATCNCNGAACNATGACAAAAGACAAATTTCTGCTGAGCGACTGCTGGATGCA 177
Db 1 ATGAGCGGATCAGAAACAATGACAAAAGACAAATTTCTGCTGAGCGACTGCTGGATGCA 60
Qy 178 GTGAAACAGTGCAGATCCGCTTTNGAGGGAGAAAGAGATTGCTCTCGATTCCGACAGC 237
Db 61 GTGAAACAGTGCAGATCCGCTTTNGAGGGAGAAAGAGATTGCTCTCGATTCCGACAGC 120
Qy 238 AGGTCACCTGTCTGTGTGCCAGTTTGAAGCGCTCTGAGCATGCTTGAAGAGAGT 297
Db 121 AGGTCACCTGTCTGTGTGCCAGTTTGAAGCGCTCTGAGCATGCTTGAAGAGAGT 180
Qy 298 CGAGGATTGCACTCAGCGCGCAGCATCAAGCAGCGAGCGGCTTTGCCAGCAAAACC 357
Db 181 CGAGGATTGCACTCAGCGCGCAGCATCAAGCAGCGAGCGGCTTTGCCAGCAAAACC 240
Qy 358 GAAACAGAGCCCGTGTCTGTGCTACTACGTCAAGAGGCTCTCAACAGCACGAGCTGCAG 417
Db 241 GAAACAGAGCCCGTGTCTGTGCTACTACGTCAAGAGGCTCTCAACAGCACGAGCTGCAG 300
Qy 418 CGCTTCTACTCTCCCTGCGCCACATCGCTCAGAGCTGGCGCGGCTCGCGCTGGCTGCGC 477
Db 301 CGCTTCTACTCTCCCTGCGCCACATCGCTCAGAGCTGGCGCGGCTCGCGCTGGCTGCGC 360
Qy 478 TGTGCTCTCAAGCAACACTCCCTGAGCGCTACCTGACATGCTCTGCGCGACCGCTGC 537
Db 361 TGTGCTCTCAAGCAACACTCCCTGAGCGCTACCTGACATGCTCTGCGCGACCGCTGC 420
Qy 538 AGGCTGAGCACTTTTATGAAGACTGGTCTTTGTGATGATGAAGAAAGCTCAGATG 597
Db 421 AGGCTGAGCACTTTTATGAAGACTGGTCTTTGTGATGATGAAGAAAGCTCAGATG 480
Qy 598 CTTCTTACCATGGCAGCAGGTCTGAACCTCACTCTTTGCGATTAACTCGACAACAG 657
Db 481 CTTCTTACCATGGCAGCAGGTCTGAACCTCACTCTTTGCGATTAACTCGACAACAG 540
Qy 658 GATTTGAACGGCGCAGAGTAAGTTTGTCTCCACCGTTTTCAGACTCTTAAAGGAGTCAACG 717
Db 541 GATTTGAACGGCGCAGAGTAAGTTTGTCTCCACCGTTTTCAGACTCTTAAAGGAGTCAACG 600
Qy 718 CAGAACGTGACCTCTCTGCTGCTGAAGAGTCCACGCAAGAGTGAAGAGCTGTTCAGGAG 777
Db 601 CAGAACGTGACCTCTCTGCTGCTGAAGAGTCCACGCAAGAGTGAAGAGCTGTTCAGGAG 660
Qy 778 ATCAGAGCTCTCTGCTGCTGCTCTCCATCTCATCAACCTGAACAGAGACCGA -CCCTTG 836
Db 661 ATCAGAGCTCTCTGCTGCTGCTCTCCATCTCATCAACCTGAACAGAGACCGA -CCCTTG 720
Qy 837 CTTGCTGCTCAGGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 896
Db 721 CTTGCTGCTCAGGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 780
Qy 897 AAGAAAGTCAACCAATAATCTCATTTGATGATGAGAAAGATGAGCAGAACTCTGGGGAC 956
Db 781 AAGAAAGTCAACCAATAATCTCATTTGATGATGAGAAAGATGAGCAGAACTCTGGGGAC 840
Qy 957 GTGTTTAAAAAGACACCTGGGCGAGAGCTCAGAGACAACTCCGACCGCTCTCT 1016
Db 841 GTGTTTAAAAAGACACCTGGGCGAGAGAGCTCAGAGACAACTCCGACCGCTCTCT 900
Qy 1017 GTCAATATCATGTCCGCTTTGAAAGCCCTTCGGGCTTAATCCCAATGGA -TCAGAGC 1075
Db 901 GTCAATATCATGTCCGCTTTGAAAGCCCTTCGGGCTTAATCCCAATGGAAGTCAGAGC 960

Db 181 CGAGATTGGCACTCAGCGGCGAGCGGATCAAGCAGCGAGCGGGCTTTGGCAGCAAAACC 240
Qy 358 GAAACAGAGCCGGTGTCTGTACTACGTGAAGAGGTCTCTCAACAGACGAGCTGCAG 417
Db 241 GAAACAGAGCCGGTGTCTGTACTACGTGAAGAGGTCTCTCAACAGACGAGCTGCAG 300
Qy 418 CGCTTCTACTCCCTGCGGCACATCGCTCAGACGTGGCCCGGGGTGCGGCTGGCTGCGC 477
Db 301 CGCTTCTACTCCCTGCGGCACATCGCTCAGACGTGGCCCGGGGTGCGGCTGGCTGCGC 360
Qy 478 TGTGCGCTCAACGAACACTCCCTGGAGCGCTACTCTGCACATGCTCTCTGGCCGACCGCTGC 537
Db 361 TGTGCGCTCAACGAACACTCCCTGGAGCGCTACTCTGCACATGCTCTCTGGCCGACCGCTGC 420
Qy 538 AGCGTGAAGCACTTTTATGAAGACTGTGTCTTTTGTGATGAAGAAAGSTCCAGTATG 597
Db 421 AGCGTGAAGCACTTTTATGAAGACTGTGTCTTTTGTGATGAAGAAAGSTCCAGTATG 480
Qy 598 CTTCTCTACATGGCAGCAGTCTGAACCTCCTATCTCTTTCGGATTAAACATCGACAAG 657
Db 481 CTTCTCTACATGGCAGCAGTCTGAACCTCCTATCTCTTTCGGATTAAACATCGACAAG 540
Qy 658 GATTGTAACGGGAGAGTAAGTTTGCTCCACCGTTTTCAGACCTCTTAAAGAGTCAACG 717
Db 541 GATTGTAACGGGAGAGTAAGTTTGCTCCACCGTTTTCAGACCTCTTAAAGAGTCAACG 600
Qy 718 CAGAACGTGACCTCTCTTGTGAGGAGTCCACGCAAGAGTGAAGCCTCTTTCAGGGAG 777
Db 601 CAGAACGTGACCTCTCTTGTGAGGAGTCCACGCAAGAGTGAAGCCTCTTTCAGGGAG 660
Qy 778 ATCAGAGCTCTCTGCGCTCTCATCTCATCAACCTGAAACAGAGACCGA-CGCTTG 836
Db 661 ATCAGAGCTCTCTGCGCTCTCATCTCATCAACCTGAAACAGAGACCGA-CGCTTG 720
Qy 837 CCGTCTGTGCCAGGATGTCAGTGTGATGCCAATGCAAAAGGCGGAGAGAA 896
Db 721 CCGTCTGTGCCAGGATGTCAGTGTGATGCCAATGCAAAAGGCGGAGAGAA 780
Qy 897 AAGAAAGTGAACCAATAATCTCATTTGATGAGGAAGATCAGCAGAACTCTGGGGAC 956
Db 781 AAGAAAGTGAACCAATAATCTCATTTGATGAGGAAGATCAGCAGAACTCTGGGGAC 840
Qy 957 GTGTTTAAAGACACCTGGGGCAGGGGAGTCAAGAGACAACTCCCGACCGCTCTCT 1016
Db 841 GTGTTTAAAGACACCTGGGGCAGGGGAGTCAAGAGACAACTCCCGACCGCTCTCT 900
Qy 1017 GTCAATATCATGTCGCGCTTTGAAGCCCTTCGGGCTTAATCTCCATGGA-TCAGAGC 1075
Db 901 GTCAATATCATGTCGCGCTTTGAAGCCCTTCGGGCTTAATCTCCATGGAAGTCAAGC 960
Qy 1076 AGCAACTCATGGAATAATTGATTCCTGTCTTTGAAAGGGGAGTTTGGGTACCAAGAGCTT 1135
Db 961 AGCAACTCATGGAATAATTGATTCCTGTCTTTGAAAGGGGAGTTTGGGTACCAAGAGCTT 1020
Qy 1136 GATGTGAAAGACATCGATGATGAAGATGTGGATGAAACGAAGATGACGTGTATGAAAC 1195
Db 1021 GATGTGAAAGACATCGATGATGAAGATGTGGATGAAACGAAGATGACGTGTATGAAAC 1080
Qy 1196 TCATCAGGAAGGAGCAGCAGGGGCCACTCGAGTCCGCCGAGAGCCACTGGAGGGAAC 1255
Db 1081 TCATCAGGAAGGAGCAGCAGGGGCCACTCGAGTCCGCCGAGAGCCACTGGAGGGAAC 1140
Qy 1256 ACCTGCTCTCCAGATGACAGCTGGGCTCCGCTGAAGTGTGACATGATGATCCGAC 1315
Db 1141 ACCTGCTCTCCAGATGACAGCTGGGCTCCGCTGAAGTGTGACATGATGATCCGAC 1200
Qy 1316 ATCCTCTCTCCCTGTCAGTGGGCTCTCTACAGCCAGCAGATGCCCTCCGGAAGC 1375
Db 1201 ATCCTCTCTCCCTGTCAGTGGGCTCTCTACAGCCAGCAGATGCCCTCCGGAAGC 1260
Qy 1376 CTGGAGAACGGGACAGCAGAGGACCAAGTCTCTCCGATCTCGAATTCGGTACAGT 1435
Db 1261 CTGGAGAACGGGACAGCAGAGGACCAAGTCTCTCCGATCTCGAATTCGGTACAGT 1320

Qy 1436 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCTCTCTGAGCAGCCTGTTA-CTTCTGCCTCA 1494
Db 1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCTCTCTGAGCAGCCTGTTACCTTCTGCCTCA 1380
Qy 1495 GTGCCAGAGTCCATGACAAATTAGTGAACCTGGCCGAGCCCACTGTGCCATGATGAACAGG 1554
Db 1381 GTGCCAGAGTCCATGACAAATTAGTGAACCTGGCCGAGCCCACTGTGCCATGATGAACAGG 1440
Qy 1555 AAGGATGAGCTGGAGGAGGAGAACAGATCACTCGAAACCTGTCTGACCGTGAATGGAG 1614
Db 1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTCGGAAACCTGTCTGACCGTGAATGGAG 1500
Qy 1615 CACTCAGCCGCTCCGGCAAGAGTGGACACCTTGAAGAGGAGTGGTGTGAACAGGAG 1674
Db 1501 CACTCAGCCGCTCCGGCAAGAGTGGACACCTTGAAGAGGAGTGGTGTGAACAGGAG 1560
Qy 1675 GAGCGCAGGGCATGAAGTCCAGGCGCTGGCCAG 1709
Db 1561 GAGCGCAGGGCATGAAGTCCAGGCGCTGGCCAG 1595

RESULT 4

AAD04334
ID AAD04334 standard; cDNA; 3139 BP.
XX
AC AAD04334;
AC AC
DT 04-JUL-2001 (first entry)
XX
DE Human TNF receptor-associated factor binding protein, IREN-10B cDNA.
XX
KW Human; Tumour Necrosis Factor; TNF; TNF Receptor Associated Factor;
KW TRAF2; TRAF2 binding protein; IkappaB Regulator; IREN-10B;
KW immunosuppressive; nuclear factor-kappaB; NF-kappaB; cytosolic; tumour;
KW AIDS; acquired immune deficiency syndrome; rheumatic disease; apoptosis;
KW autoimmune disease; septic shock; graft-vs-host reaction; inflammation;
KW anorexia; anti-HIV; therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..2442
FT /*tag= a
FT /product= "Human TRAF2-binding protein isoform, IREN-10B"
XX
PN WO200116314-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-IL000517.
XX
PR 02-SEP-1999; 99IL-00131719.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Wallach D, Malinin N, Sinha I, Leu S;
XX
DR WPI; 2001-281387/29.
DR P-PSDB; AAE00684.

XX
PT New DNA sequence encoding Tumor Necrosis Factor receptor associated
PT factor (TRAF) binding proteins (IREN) for treatment or prevention of
PT pathological conditions associated with NF-kappaB induction.
XX
PS Claim 1; Fig 4; 118pp; English.

XX
CC The present cDNA sequence encodes human tumour necrosis factor (TNF)
CC receptor-associated factor (TRAF2) binding protein isoform, IREN-10B. A
CC fragment of this IREN-10B molecule is capable of binding to TRAF2 protein
CC at position 225-501. The invention relates to human tumour necrosis
CC factor (TNF) receptor-associated factor (TRAF2) binding protein
CC designated as IREN (IkappaB Regulator), its isoforms IREN-10B, IREN-E and

CC their corresponding cDNA molecules. IREN is useful for
 CC modulating/mediating the activity of transcription factor NF (Nuclear
 CC Factor)-kappaB or any other intracellular signalling activity mediated by
 CC TRAF2. IREN is useful in the prevention and treatment of a pathological
 CC condition associated with NF-kappaB induction (abnormal) e.g. AIDS
 CC (acquired immune deficiency syndrome), autoimmune diseases, tumours,
 CC rheumatic diseases, anorexia, septic shock and graft-vs-host reactions.
 CC IREN also plays an important role in the control of inflammation and
 CC other non-apoptotic effects of TNF as well as in the control of
 CC apoptosis. The invention also relates to method for screening,
 CC identifying and producing a molecule capable of modulating activities
 CC mediated by IREN. IREN antibodies are useful for the purification of new
 CC proteins from different sources, including cell extracts or transformed
 CC cell lines, in addition IREN can be used in diagnostic purposes for
 CC identifying disorders related to abnormal functioning of cellular effects
 CC mediated directly by TRAF proteins
 XX

SQ Sequence 3139 BP; 797 A; 851 C; 875 G; 616 T; 0 U; 0 Other;

Query Match 81.6%; Score 1555; DB 4; Length 3139;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1588; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

QY 118 ATGCGCGATCNCGAACATACAAAGACAAATTTCTGCGAGCGACTGCTGATGCA 177
 DB 1 ATGAGCGGATCACGAACAATGACAAAAGACAAATTTCTGCTGGAGCGACTGCTGATGCA 60
 QY 178 GTGAACAGTGCAGATCCGCTTTNGAGGAGAGAGAGATTGCTCCGATTCCGACAGC 237
 DB 61 GTGAACAGTGCAGATCCGCTTTGAGGAGAGAGAGATTGCTCCGATTCCGACAGC 120
 QY 238 AGGTCACCTGTCGTGTCGCCAGTTTGAAGCGTCTGCGACATGGCTTGAAGAGAGT 297
 DB 121 AGGTCACCTGTCGTGTCGCCAGTTTGAAGCGTCTGCGACATGGCTTGAAGAGAGT 180
 QY 298 CGAGGATTGGCACTCACAGCGCGACGATCAGCAGCGAGCGGCTTTGCCAGCAAAACC 357
 DB 181 CGAGGATTGGCACTCACAGCGCGACGATCAGCAGCGAGCGGCTTTGCCAGCAAAACC 240
 QY 358 GAAACAGAGCCGTTTCTGCTGCTACTACGTGAAGGAGTCTTCAACAGCAGCGATGCGAG 417
 DB 241 GAAACAGAGCCGTTTCTGCTGCTACTACGTGAAGGAGTCTTCAACAGCAGCGATGCGAG 300
 QY 418 CGCTTCTACTCCCTCGCGCCATCGCTCAGACGTGGCCGGGTCGCGCTGCTGCGC 477
 DB 301 CGCTTCTACTCCCTCGCGCCATCGCTCAGACGTGGCCGGGTCGCGCTGCTGCGC 360
 QY 478 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 537
 DB 361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGC 420
 QY 538 AGGCTGAGCATTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGTCCAGTATG 597
 DB 421 AGGCTGAGCATTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGTCCAGTATG 480
 QY 598 CTTCTACCATGGCAGCAGGTCTGAATCCATCTCTTTGCGATTAACATCGACAAACAG 657
 DB 481 CTTCTACCATGGCAGCAGGTCTGAATCCATCTCTTTGCGATTAACATCGACAAACAG 540
 QY 658 GATTGGAACGGGACAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACG 717
 DB 541 GATTGGAACGGGACAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACG 600
 QY 718 CAGAACTGACCTCTTGTGAGGAGTCCACGCAAGGAGTGACAGCCTGTTTCAGGGAG 777
 DB 601 CAGAACTGACCTCTTGTGAGGAGTCCACGCAAGGAGTGACAGCCTGTTTCAGGGAG 660
 QY 778 ATCAGCCTCTCTGCGCTCTCCATCTCATCAAACTCAACAGGAGACCGA-CCCTTG 836
 DB 661 ATCAGCCTCTCTGCGCTCTCCATCTCATCAAACTCAACAGGAGACCGA-CCCTTG 720
 QY 837 CCTGCTGTGCCAGAAATGTGCTGCTGATGCCAAATGCAAAAGAGCGGAGAAAGAAA 896
 |||

DB 721 CCTGCTGTCAGGAATGTGCTGCTGATGCCAAATGCAAAAGGAGCGGAGAAAGAAA 780
 QY 897 AAGAAAGTGACCAACAATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 956
 DB 781 AAGAAAGTGACCAACAATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840
 QY 957 GTGTTTAAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTCT 1016
 DB 841 GTGTTTAAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTCT 900
 QY 1017 GTCAATATCATGTCGGCTTTTAAAGGCCCTTTCGGGCTTAACCTCAATGGAA-TCAGAGC 1075
 DB 901 GTCAATATCATGTCGGCTTTTAAAGGCCCTTTCGGGCTTAACCTCAATGGAAAGTCAGAGC 960
 QY 1076 AGCAACTCATGGAATAATGATTCCTGTCTTTGAACGGGAGTTTGGGTACCAAGAGCTT 1135
 DB 961 AGCAACTCATGGAATAATGATTCCTGTCTTTGAACGGGAGTTTGGGTACCAAGAGCTT 1020
 QY 1136 GATGTGAAAAGCATCGATGATGAAGATGTGATGAAAACGAAAGATGACGTGTATGGAAC 1195
 DB 1021 GATGTGAAAAGCATCGATGATGAAGATGTGATGAAAACGAAAGATGACGTGTATGGAAC 1080
 QY 1196 TCATCAGAAAGGAAGACACAGGGGCCACTCGGAGTTCGCCGAGAGCCACTGGAAGGGAAC 1255
 DB 1081 TCATCAGAAAGGAAGACACAGGGGCCACTCGGAGTTCGCCGAGAGCCACTGGAAGGGAAC 1140
 QY 1256 ACCTGCTCTCCAGATGCAAGCTGGGCTCCGCTGAGGTGCTGCAATGATCTCGAC 1315
 DB 1141 ACCTGCTCTCCAGATGCAAGCTGGGCTCCGCTGAGGTGCTGCAATGATCTCGAC 1200
 QY 1316 ATCTCTTTCCTGTGCTGCTGGGCTCTTACAGCCAGCAGATGCCCCCTCGGAAGC 1375
 DB 1201 ATCTCTTTCCTGTGCTGCTGGGCTCTTACAGCCAGCAGATGCCCCCTCGGAAGC 1260
 QY 1376 CTGAGAAACGGGACAGGACACAGAGGACCAAGTTTCTCCCGATCTCGGACTTCGGTACAGT 1435
 DB 1261 CTGAGAAACGGGACAGGACACAGAGGACCAAGTTTCTCCCGATCTCGGACTTCGGTACAGT 1320
 QY 1436 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCTCTGAGCAGCGCTGTA-CTTCTGCTCA 1494
 DB 1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCTCTGAGCAGCGCTGTA-CTTCTGCTCA 1380
 QY 1495 GTGCCAGAGTCCATGACAAATTAGTGAATCTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1554
 DB 1381 GTGCCAGAGTCCATGACAAATTAGTGAATCTGCGCCAGGCCACTGTGGCCATGATGAACAGG 1440
 QY 1555 AAGGATGAGCTGAGGAGGAGAAACAGATCACTCGGAAAACCTGCTCGACGGTGAAGTGGAG 1614
 DB 1441 AAGGATGAGCTGAGGAGGAGAAACAGATCACTCGGAAAACCTGCTCGACGGTGAAGTGGAG 1500
 QY 1615 CACTCAGCCGCTCCCGCAAGAGTGGACACCTTCAAAAGGAGGTGGCTGAACAGGAG 1674
 DB 1501 CACTCAGCCGCTCCCGCAAGAGTGGACACCTTCAAAAGGAGGTGGCTGAACAGGAG 1560
 QY 1675 GAGCGCAGGGCATGAAGGTCCAGGCGCTGGCCAG 1709
 DB 1561 GAGCGCAGGGCATGAAGGTCCAGGCGCTGGCCAG 1595

RESULT 5

AAH15665

ID AAH15665 standard; cDNA; 2248 BP.

XX

XX AAH15665;

XX AC

DT 26-JUN-2001 (first entry)

XX

DE Human cDNA sequence SEQ ID NO:14016.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX

OS Homo sapiens.

XX

PR	14-AUG-2000;	2000US-0225270P.
PR	14-AUG-2000;	2000US-0225447P.
PR	14-AUG-2000;	2000US-0225757P.
PR	14-AUG-2000;	2000US-0225758P.
PR	14-AUG-2000;	2000US-0225759P.
PR	14-AUG-2000;	2000US-0226279P.
PR	14-AUG-2000;	2000US-0226681P.
PR	14-AUG-2000;	2000US-0226686P.
PR	22-AUG-2000;	2000US-0227182P.
PR	23-AUG-2000;	2000US-0227009P.
PR	30-AUG-2000;	2000US-0228924P.
PR	01-SEP-2000;	2000US-0228287P.
PR	01-SEP-2000;	2000US-0229343P.
PR	01-SEP-2000;	2000US-0229344P.
PR	05-SEP-2000;	2000US-0229509P.
PR	05-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0230437P.
PR	06-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231124P.
PR	08-SEP-2000;	2000US-0231243P.
PR	08-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231413P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0232060P.
PR	08-SEP-2000;	2000US-0232081P.
PR	13-SEP-2000;	2000US-0231968P.
PR	14-SEP-2000;	2000US-0232397P.
PR	14-SEP-2000;	2000US-0232398P.
PR	14-SEP-2000;	2000US-0232399P.
PR	14-SEP-2000;	2000US-0232400P.
PR	14-SEP-2000;	2000US-0232401P.
PR	14-SEP-2000;	2000US-0233063P.
PR	14-SEP-2000;	2000US-0233064P.
PR	14-SEP-2000;	2000US-0233065P.
PR	21-SEP-2000;	2000US-0234223P.
PR	21-SEP-2000;	2000US-0234274P.
PR	25-SEP-2000;	2000US-0234997P.
PR	25-SEP-2000;	2000US-0234998P.
PR	26-SEP-2000;	2000US-0235484P.
PR	27-SEP-2000;	2000US-0235834P.
PR	27-SEP-2000;	2000US-0235836P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.
PR	29-SEP-2000;	2000US-0236368P.
PR	29-SEP-2000;	2000US-0236369P.
PR	02-OCT-2000;	2000US-0236370P.
PR	02-OCT-2000;	2000US-0236802P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	13-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239935P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241221P.
PR	20-OCT-2000;	2000US-0241783P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0241826P.
PR	01-NOV-2000;	2000US-0244617P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246529P.
PR	08-NOV-2000;	2000US-0246530P.
PR	08-NOV-2000;	2000US-0246531P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246533P.
PR	08-NOV-2000;	2000US-0246534P.
PR	08-NOV-2000;	2000US-0246535P.
PR	08-NOV-2000;	2000US-0246536P.
PR	08-NOV-2000;	2000US-0246537P.
PR	08-NOV-2000;	2000US-0246538P.
PR	08-NOV-2000;	2000US-0246539P.
PR	08-NOV-2000;	2000US-0246540P.
PR	08-NOV-2000;	2000US-0246541P.
PR	08-NOV-2000;	2000US-0246542P.
PR	08-NOV-2000;	2000US-0246543P.
PR	08-NOV-2000;	2000US-0246544P.
PR	08-NOV-2000;	2000US-0246545P.
PR	08-NOV-2000;	2000US-0246546P.
PR	08-NOV-2000;	2000US-0246547P.
PR	08-NOV-2000;	2000US-0246548P.
PR	08-NOV-2000;	2000US-0246549P.
PR	08-NOV-	

Qy	862	CTGATGCCAAATGCAAAAAGAGCGGGAAGAAGAAAAAGAAAGTCACCAACATAATCTCAT	921
Db	8894	CAGATGCCAAATGCAAAAAGAGCGGGAAGAAGAAAAAGAAAGTCACCAACATAATCTCAT	8953
Qy	922	TTGATGATGAGGAAGATGAGCAGAACTCTCGGGACGTGTTTAAAAAGACACCTCGGGCGAG	981
Db	8954	TTGATGATGAGGAAGATGAGCAGAACTCTCGGGACATGTTTAAAAAGACACCTCGGGCGAG	9013
Qy	982	GGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAATATCATATGTCGCGCTTTGAAA	1041
Db	9014	GGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAATATCATATGTCGCGCTTTGAAA	9073
Qy	1042	GCCCTTTGGGCGCTTAATCCAATGGAA-TCAGACGACGCAACTATGAAAAATTTGATTTCCC	1100
Db	9074	GCCCTTTGGGCGCAAACTCCAATGGAAGTCAGACGACGCAACTCGTGGAAAAATTTGATTTCCC	9133
Qy	1101	TGCTTTTGAACGGGGAGTTTGGGTACAGAAAGCTTGATGTGAAAAAGCATCGATGATGAAG	1160
Db	9134	TGCTTTTGAACAGGGAGTTTGGGTACAGAAAGCTTGATGTGAAAAAGCATCGATGATGAAG	9193
Qy	1161	ATGTGGATGAAAACGAAGATGACGTGTATGGAACCTCATCAGGAAGGAAGCAGACGGGGCC	1220
Db	9194	ATGTGGATGAAAACGAAGATGACGTGTATGGAACCTCATCAGGACGGAAGCAGACGGGGCC	9253
Qy	1221	ACTCGGAGTCGCCCCGAGAAGCCACT	1245
Db	9254	ACTCAGAGTCGCCCCGAGAAGTAACT	9278

RESULT 9

AAS64410

ID AAS64410 standard: cDNA: 454 BP.

XX
XX

AC AAS64410;

XX

DT	13-FEB-2002 (first entry)
XX	
XX	DNA encoding novel human diagnostic protein #214.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens

XX
XX

PN WO200175067-A2.

[illegible]FD 11-OCT-2001.
XXXX
PF 30-MAR-2001: 2001WO-IIS008631XX
XX
OMT007; T007-WYJ-06

PR 31-MAR-2000: 2000US-

PR	23-AUG-2000; 2000US-00649167.
XX	(HYSE-) HYSEQ INC.
PA	Dmanac RT, Liu C, Tang YT;
PI	
XX	WPI; 2001-639362/73.
DR	P-PSDB; ABG00223.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.

XX
PS
XX
Claim 1; SEQ ID NO 214; 103pp; English.
XX
The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed

genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/publ/published_pct_sequences](http://wipo.int/pub/publ/published_pct_sequences)

XX

Query Match

Best Local Similarity 96.2%; Pred. No. 3.3e-64;

Matches 277; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 77 GCCGCGCGCGCGCGGANGCACC GCCCGCGGGANAGCNCATGACCGGATCNCNGAACN 136

Db 27 GCGGCGCGGTGCAGGCA^{CCGTTCCGGGAGAGCACCATGAGCGGATCACAGAACA} 86

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY
137 ATGACAAAAGACAAATTCTGCTGGAGCGACTGCTGGATGCAGTGAAACAGTGCCAGATCC 196
QY

Db 87 ATGACAAAAGACAAATTTCTGCTGGAGCGACTGCTGGATGCAGTGAACACAGTGCACATCC 146

DD 8 / ATGACAAATAGACAAATTTCTGCTGGAGCGACCTGCTGGATGAGTATGACAGATGCTCAGATCC 148

QY 197 GCTTNGAGGGAGAAAGGAGATTGCCCTCGGATTCCGACAGCAGGGTCACCTGCTGTGTG 256

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

RESULT

RESOL TO
ABS72233

ID ABS72233 standard; cDNA; 280 BP.

XX

AC ABS72233;

XX

DT	03-DEC-2002	(first entry)	
XX			Human gene trapped sequence (GTS) #193.
DE			
XX			
KW			Human; gene trapped sequence; GTS; gene; se; cancer; autoimmune disease; lupus; scleroderma; Crohn's disease; multiple sclerosis; immune disorder
KW			inflammatory bowel disease; schizophrenia; psychosis; osteoarthritis;
KW			inflammatory disorder; diabetes; skin disorder; acne; eczema; asthma;
KW			rheumatoid arthritis; hypertension; atherosclerosis; Alzheimer's disease
KW			cardiovascular disease; Parkinson's disease; osteoporosis; infertility;
KW			viral infection; parasitic infection; fungal infection;
KW			bacterial infection; forensic analysis; cellular differentiation.

44
44
45
44
43
40
40
45
42

OS Homo sapiens.

XXIV

PN US2002095031-A1.

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	

PD

PA (NEHL/) NEHLS M C.
PA (ZAMB/) ZAMBROWICZ B.
PA (SAND/) SANDS A T.
XX
PI Nehls MC, Zambrowicz B, Sands AT;
XX WPI; 2002-656030/70.
XX
XX New isolated or purified human gene trapped sequences, useful for gene
PT discovery, as markers for gene expression analysis, identifying and
PT mapping the coding regions of human genome, or determining the genetic
PT basis of human disease.
XX
XX Claim 1; SEQ ID NO 201; 36pp; English.
PS
XX The invention relates to isolated or purified polynucleotides that
CC correspond to human gene trapped sequences (GTSs). The human GTSs are
CC useful for gene discovery and as markers for gene expression analysis,
CC for identifying and mapping the coding regions of the mammalian,
CC particularly human, genome, for forensic analysis, and for determining
CC the genetic basis of human disease. The peptides and proteins encoded by
CC the polynucleotides are useful for generating antibodies, as reagents in
CC diagnostic assays and in identifying other cellular gene products
CC involved in the regulation of development and cellular differentiation of
CC various cell types. The peptides are also useful as reagents in assays
CC for screening of compounds used in treating disorders affecting
CC development and cell differentiation. The GTSs are also useful in
CC treating or ameliorating diseases associated with the expression of
CC mutant or normal variants of the GTSs, e.g. cancer, autoimmune diseases,
CC lupus, scleroderma, Crohn's disease, multiple sclerosis, inflammatory
CC bowel disease, immune disorders, schizophrenia, psychosis, inflammatory
CC disorders, diabetes, skin disorders such as acne or eczema,
CC osteoarthritis, rheumatoid arthritis, hypertension, atherosclerosis,
CC cardiovascular diseases, Alzheimer's disease, Parkinson's disease,
CC osteoporosis, asthma, infertility, and viral, parasitic, fungal or
CC bacterial infections. This sequence represents a human GTS of the
XX invention
XX
SQ Sequence 280 BP; 71 A; 66 C; 91 G; 51 T; 0 U; 1 Other;
Query Match 13.8%; Score 263.8; DB 6; Length 280;
Best Local Similarity 98.6%; Pred. No. 2.3e-61;
Matches 276; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1464 GTCTCTGAGCGCTGTGA-CTTCTGCTCAGTCCAGAGTCCATGACATTAGTGAAC 1522
DB 1 GGCCTATGAGCGCCTGNTACCTTCTGCTCAGTCCAGAGTCCATGACATTAGTGAAC 60
QY 1523 TGGCCAGGCCACTGTGGCCATGATCAACAGGAAGGATGAGCTGGAGGAGAGACAGAT 1582
DB 61 TGGCCAGGCCACTGTGGCCATGATCAACAGGAAGGATGAGCTGGAGGAGAGACAGAT 120
QY 1583 CACTCGAAACCTGTCTGACGGTGAGATGAGACATCTAGCCCGCTCCGGCAAGAGGTGG 1642
DB 121 CACTCGAAACCTGTCTGACGGTGAGATGAGACATCTAGCCCGCTCCGGCAAGAGGTGG 180
QY 1643 ACACCTTGAAGGAGGTGGTGTACAGAGAGCGGCGAGGCGATGAGTCCAGGCGC 1702
DB 181 ACACCTTGAAGGAGGTGGTGTACAGAGAGCGGCGAGGCGATGAGTCCAGGCGC 240
QY 1703 TGGCCAGCTATCTTTGCTATTTTGTGAGGAGATTTCTAAC 1742
DB 241 TGGCCAGCTATCTTTGCTATTTTGTGAGGAGATTTCTAAC 280
RESULT 11
ID ACH75347/c
XX ACH75347 standard; DNA; 542 BP.
XX ACH75347;
XX 29-JUL-2004 (first entry)
XX

DE Human genome derived single exon probe #8542.
XX Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
XX Homo sapiens.
OS
XX US2003194704-A1.
PN
XX 16-OCT-2003.
PD
XX 03-APR-2002; 2002US-00029386.
PF
XX 03-APR-2002; 2002US-00029386.
PR
XX (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANK/) HANKEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
XX WPI; 2004-119264/12.
DR
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
PT
XX Claim 15; SEQ ID NO 8542; 80pp; English.
PS
XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 542 BP; 135 A; 143 C; 129 G; 135 T; 0 U; 0 Other;
Query Match 11.3%; Score 215.8; DB 12; Length 542;
Best Local Similarity 95.7%; Pred. No. 4e-48;
Matches 244; Conservative 0; Mismatches 7; Indels 4; Gaps 2;
QY 614 CAGGTCTGAATCCATCTCTTTGCGATTACATCGACACAGGATTGTAACGGGCAGA 673

Db 495 CAGTCCGAACTCATCTCTTTCGGATTAACTGACAAAGGATTGAACGGGCGA 436
Qy 674 GTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACCGCAGAACGTGACCTCT 733
Db 435 GTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACCGCAGAACGTG 379
Qy 734 TGCTGAAGGAGTCCACGCAAGGAGTGAGCGCTGTTTCAGGAGATCACAGCTCTCTCTG 793
Db 378 TGCTGAAGGAGTCCACGCAAGGAGTGAGCGCTGTTTCAGGAGATCACAGCTCTCTCTG 319
Qy 794 CCGTCTCCATCCCTCATCAACCTGAACAGGAGACCGA-CCTTGCCTGTGCTGCCAGGA 852
Db 318 CCATCTCCATCTCATCAACCTGAACAGGAGACCGCTTGCCTGTGCTGCCAGGA 259
Qy 853 ATGTCAGTGTGTG 867
Db 258 ATGTCAGTGTGTG 244

RESULT 12
ACH89047/c
ID ACH89047 standard; DNA; 284 BP.
AC ACH89047;
XX
DT 29-JUL-2004 (first entry)
XX Human genome derived single exon probe #22242.
XX Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX Homo sapiens.
XX US2003194704-A1.
XX 16-OCT-2003.
XX 03-APR-2002; 2002US-00029386.
XX 03-APR-2002; 2002US-00029386.
XX (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX Penn SG, Rank DR, Hanzel DK;
XX WPI; 2004-119264/12.
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX Claim 1; SEQ ID NO 22242; 80pp; English.
XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid

CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subexpression, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 284 BP; 60 A; 78 C; 74 G; 72 T; 0 U; 0 Other;

Query Match 11.2%; Score 212.8; DB 12; Length 284;
Best Local Similarity 95.8%; Pred. No. 1.9e-47;
Matches 241; Conservative 0; Mismatches 7; Indels 4; Gaps 2;
Qy 617 GTCTGAATCCATCTCTTTGCGATTAACTGACAAAGGATTGAACGGGCGAGTA 676
Db 284 GTCCGAATCCATCTCTTTGCGATTAACTGACAAAGGATTGAACGGGCGAGTA 225
Qy 677 AGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACCGCAGAACGTGACCTCTCTG 736
Db 224 AGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACCGCAGAACGTG 168
Qy 737 TGAAGGAGTCCACGCAAGGAGTGAGCGCTGTTCAGGGAGATCACAGCTCTCTCTCCG 796
Db 167 TGAAGGAGTCCACGCAAGGAGTGAGCGCTGTTCAGGGAGATCACAGCTCTCTCTCC 108
Qy 797 TCTCCATCTCTCATCAACCTGAACAGGAGACCGA-CCTTGCCTGTGCTGCCAGGAATG 855
Db 107 TCTCCATCTCTCATCAACCTGAACAGGAGACCGA-CCTTGCCTGTGCTGCCAGGA 48
Qy 856 TCAGTGTCTGATG 867
Db 47 TCAGTGTCTG 36

RESULT 13
AAS92200
ID AAS92200 standard; cDNA; 463 BP.
XX
AC AAS92200;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #28004.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX W0200175067-A2.
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US0008631.
XX
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.

```
XX PI Drmanac RT, Liu C, Tang YT;
XX PD WPI: 2001-639362/73.
XX PF P-PSDB; ABG28013.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 1; SEQ ID NO 28004; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX CC coding sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 463 BP; 124 A; 133 C; 122 G; 84 T; 0 U; 0 Other;
XX Query Match 10.3%; Score 195.6; DB 5; Length 463;
XX Best Local Similarity 93.5%; Pred. No. 1.2e-42;
XX Matches 215; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
XX QY 1334 GGCTGGGCTCTACAGCCCGAGATGCCCTCGGAAGCCTGGAGAACGGACAGGA 1393
XX Db 234 GTCTCAGCCCTGGCTGAACAAAGATGCCCTCGGAAGCCTGGAGAACGGACAGGA 293
XX QY 1394 CCAGAGGACCGCTTCTCCGGATCCTGGATTCGGTACAGTGTGGAGCCAGCTCTCCA 1453
XX Db 294 CCAGAGGACCGCTTCTCCGGATCCTGGATTCGGTACAGTGTGGAGCCAGCTCTCCA 353
XX QY 1454 GGCCACGGAAGTCTCTCAGCAGCCTGTTA-CTTCTGCTCAGTGCCAGAGTCCATGACA 1512
XX Db 354 GGCCACGGAAGTCTCTCAGCAGCCTGTTA-CTTCTGCTCAGTGCCAGAGTCCATGACA 413
XX QY 1513 ATTAGTGAATGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGA 1562
XX Db 414 ATTAGTGAATGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGA 463
XX RESULT 14
XX AAS73549/c
XX ID AAS73549 standard; cDNA; 474 BP.
XX XX AAS73549;
XX AC AAS73549;
XX XX 13-FEB-2002 (first entry)
XX DT DNA encoding novel human diagnostic protein #9353.
XX DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX XX
```

```
FN WO200175067-A2.
XX XX 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX XX (HYSE-) HYSEQ INC.
XX PA Drmanac RT, Liu C, Tang YT;
XX PI WPI: 2001-639362/73.
XX DR P-PSDB; ABG03362.
XX CC New isolated polynucleotide and encoded polypeptides, useful in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits and to assess
XX CC biodiversity.
XX PS Claim 1; SEQ ID NO 9353; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX CC coding sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 474 BP; 75 A; 143 C; 122 G; 134 T; 0 U; 0 Other;
XX Query Match 7.1%; Score 135.4; DB 5; Length 474;
XX Best Local Similarity 79.6%; Pred. No. 3.2e-26;
XX Matches 160; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
XX QY 238 AGGTCACCTGTCTGTGTGCCAGTTTGAAGCCGCTCTGCAGCATGCTTGAAGAGGAGT 297
XX Db 327 AGGTCACCTGTCTGTGTGCCAGTTTGAAGCCGCTCTGCAGCATGCTTGAAGAGGAGT 268
XX QY 298 CGAGGATTCACCTCAGCGGCGAGCGCATCAAGCAGCAGCGGCTTTGCCAGCAAAACC 357
XX Db 267 CGAGGATTCACCTCAGCGGCGAGCGCATCAAGCAGCAGCGGCTTTGCCAGCAAAACC 208
XX QY 358 GAAACAGAGCCCGTGTCTTGGTACTACGTGAAGAGAGTCTCTCAACAAGCAGCAGCTGCAG 417
XX Db 207 GAAACAGGTGAACCTGCCAGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGAGCTGGAG 148
XX QY 418 CGCTTCTACTCCCTGCCCCAC 438
XX Db 147 GAGGAGAACAGCCTGGGTGAC 127
XX RESULT 15
XX AAC11927
XX ID AAC11927 standard; cDNA; 190 BP.
XX XX AAC11927;
XX AC AAC11927;
```

XX	06-OCT-2000	(first entry)
DT	Human secreted protein 5' EST, SEQ ID NO: 16002.	
XX		
DE	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation; gene therapy; chromosome mapping; ss.	
KW	Homo sapiens.	
OS	EPI033401-A2.	
PN	06-SEP-2000.	
XX		
PD	21-FEB-2000; 2000EP-00200610.	
XX		
PF	26-FEB-1999; 99US-0122487P.	
XX		
PR	(GEST) GENSET.	
XX		
PA	Dumas Milne Edwards J, Duclert A, Giordano J;	
XX	WPI; 2000-500381/45.	
DR		
XX	New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures.	
PT		
PT	Claim 1; SEQ ID NO 16002; 71pp + Sequence Listing; English.	
XX		
PS	The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. No ORF has yet been conclusively identified within the present sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors	
CC		
CC	Sequence 190 BP; 47 A; 42 C; 73 G; 28 T; 0 U; 0 Other;	
CC		
CC	Query Match 7.0%; Score 133.4; DB 3; Length 190;	
CC	Best Local Similarity 99.3%; Pred.No.7.2e-26;	
CC	Matches 134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1578 CAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGGCTCCGGCAAGA 1637	
DB	19 CAGATCACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGGCTCCGGCAAGA 78	
QY	1638 GGTGGACACCTTGAAAAGGAAGGTGGTGCTGAACAGGAGGAGCGGCAGGCATGAAGGTCCA 1697	
DB	79 GGTGGACACCTTGAAAAGGAAGGTGGTGCTGAACAGGAGGAGCGGCAGGCATGAAGGTCCA 138	
QY	1698 GGGCGCTGGCCAGCTA 1712	
DB	139 GGGCGCTGGCCAGGTA 153	

Search completed: June 10, 2005, 06:10:07
Job time : 655.812 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 14:23:14 ; Search time 724.339 Seconds
(without alignments)
16311.995 Million cell updates/sec

Title: US-09-155-676B-1
Perfect score: 1906
Sequence: 1 cattgggtcacggtggtggc.....tcattccttgataaaaaaa 1906

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/prodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/prodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/prodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/prodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/prodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/prodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	263.8	13.8	280	9	US-09-563-817-201
2	215.8	11.3	542	16	US-10-029-386-8542
3	212.8	11.2	284	16	US-10-029-386-22242
4	128	6.7	500	16	US-10-029-386-8336
5	124	6.5	124	16	US-10-029-386-22036
6	63.8	3.3	3931	18	US-10-276-774-948
7	59.6	3.1	263	9	US-09-867-701-9696

8	54.6	2.9	371	10	US-09-918-995-24284	Sequence 24284, A
c	54.6	2.9	446	9	US-09-880-107-1518	Sequence 1518, Ap
9	54	2.8	464	10	US-09-918-995-15177	Sequence 15177, A
11	53.2	2.8	476	21	US-09-854-867-118	Sequence 118, App
12	53.2	2.8	476	21	US-10-786-970A-118	Sequence 118, App
13	50.4	2.6	7093	15	US-10-101-510-527	Sequence 527, App
14	50.4	2.6	10126	17	US-10-242-355-1024	Sequence 1024, Ap
15	50.2	2.6	550	19	US-10-437-963-63143	Sequence 63143, A
16	49.6	2.6	7140	21	US-10-211-028-3	Sequence 3, Appl
17	49.6	2.6	90597	21	US-10-211-028-1	Sequence 1, Appl
18	49.2	2.6	18535	9	US-09-764-878-385	Sequence 385, App
19	49.2	2.6	18535	14	US-10-079-854-385	Sequence 385, App
20	48	2.5	2463	13	US-10-027-632-102953	Sequence 102953,
21	48	2.5	2463	17	US-10-027-632-102953	Sequence 102953,
22	46.6	2.4	1182	13	US-10-027-348-1	Sequence 1, Appl
23	46.6	2.4	1185	13	US-10-027-348-13	Sequence 13, Appl
24	45.8	2.4	441	15	US-10-106-698-2578	Sequence 2578, Ap
25	45.6	2.4	962	20	US-10-723-860-6507	Sequence 6507, Ap
26	45	2.4	2033	9	US-09-764-868-623	Sequence 623, App
c	44.8	2.4	2056	17	US-10-151-469-13	Sequence 13, App
28	44.8	2.4	2056	19	US-10-717-597-106	Sequence 106, App
c	44.8	2.4	2056	19	US-10-802-432-14	Sequence 14, Appl
30	44.8	2.4	2056	20	US-10-723-860-2201	Sequence 2201, Ap
c	44.4	2.3	505	13	US-10-027-632-276057	Sequence 276057,
c	44.4	2.3	505	17	US-10-027-632-276057	Sequence 276057,
33	44	2.3	520	19	US-10-437-963-24663	Sequence 24663, A
34	44	2.3	550	9	US-09-925-298-379	Sequence 379, App
35	44	2.3	550	14	US-10-102-806-379	Sequence 379, App
c	44	2.3	4216	19	US-10-437-963-53652	Sequence 53652, A
c	43.8	2.3	1599	19	US-10-437-963-2771	Sequence 2771, Ap
38	43.6	2.3	531	9	US-09-925-298-381	Sequence 381, App
39	43.6	2.3	531	14	US-10-102-806-381	Sequence 381, App
40	43.4	2.3	1162	14	US-10-198-846-8983	Sequence 8983, Ap
c	43.2	2.3	505	13	US-10-027-632-276059	Sequence 276059,
c	43.2	2.3	505	17	US-10-027-632-276059	Sequence 276059,
c	43	2.3	850	18	US-10-425-114-29073	Sequence 29073, A
c	43	2.3	862	20	US-10-425-115-10220	Sequence 10220, A
45	43	2.3	940	9	US-09-789-561-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1

US-09-563-817-201
; Sequence 201, Application US/09563817
; Patent No. US20020095031A1
; GENERAL INFORMATION:
; APPLICANT: Nehls, Michael C.
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020095031A1 Human Polynucleotides and the
; FILE REFERENCE: LEX-0021-USA
; CURRENT APPLICATION NUMBER: US/09/563,817
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 60/132,343
; PRIOR FILING DATE: 1999-05-04
; NUMBER OF SEQ ID NOS: 1008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201
; LENGTH: 280
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(280)
; OTHER INFORMATION: n = A,T,C or G
US-09-563-817-201

Query Match 13.8%; Score 263.8; DB 9; Length 280;
Best Local Similarity 98.6%; Pred. No. 7.9e-71;
Matches 276; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1464 GTCCCTGAGCAGCTGTTA-CTTCTGCTCAGTGCAGAGTCCATGACAAATTAGTGAAC 1522
Db 1 GGCCATATGAGCAGCCTGNTACCTTCTGCTCAGTGCAGAGTCCATGACAAATTAGTGAAC 60
Qy 1523 TGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGAGCTGGAGGAGGAGAACAGAT 1582
Db 61 TGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGAGCTGGAGGAGGAGAACAGAT 120
Qy 1583 CACTGCGAAACCTCTCGACGGTGCAGATGGAGCACTCAGCGCGCTCCGGCAAGAGGTGG 1642
Db 121 CACTGCGAAACCTCTCGACGGTGCAGATGGAGCACTCAGCGCGCTCCGGCAAGAGGTGG 180
Qy 1643 ACACCTTGAAAGAAAGTGGCTGAACAGGAGGAGCGGCGAGGCGATGAAGTCCAGGCGC 1702
Db 181 ACACCTTGAAAGAAAGTGGCTGAACAGGAGGAGCGGCGAGGCGATGAAGTCCAGGCGC 240
Qy 1703 TGGCCAGCTATCTTTGCTATTTTGTGAGGAGATTCTAACC 1742
Db 241 TGGCCAGCTATCTTTGCTATTTTGTGAGGAGATTCTAACC 280

RESULT 2
US-10-029-386-8542/c
; Sequence 8542, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8542
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR16.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: NT HIT: g114779172, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: 014197, EVALUE 5.90e-01
; OTHER INFORMATION: EST_HUMAN HIT: BG340557.1, EVALUE 0.00e+00
US-10-029-386-8542

Query Match 11.3%; Score 215.8; DB 16; Length 542;
Best Local Similarity 95.7%; Pred. No. 8.5e-56;
Matches 244; Conservative 0; Mismatches 7; Indels 4; Gaps 2;
Qy 614 CAGGTCTGAACTCCATCTCTTTGCGATTAAACATGACAAAGGATTTGAACGGGCAGA 673
Db 495 CAGGTCCGAACCTCATACTCTTTGCGATTAAACATGACAAAGGATTTGAACGGGCAGA 436
Qy 674 GTAAGTTGCTCCACCGTTTCAGACTCTTAAAGGAGTCAACGCGACAGTGAACCTCTCT 733
Db 435 GTAAGTTGCTCCACCGTTTCAGACTCTTAAAGGAGTCAACGCGACAGTGAACCTCTCT 379
Qy 734 TGCTGAAGGAGTCCACCAAGGAGTGAAGCCTGTTTCAAGGAGATCAAGCCTCTCTCTG 793
Db 378 TGCTGAAGGAGTCCACCAAGGAGTGAAGCCTGTTTCAAGGAGATCAAGCCTCTCTCTG 319
Qy 794 CCGTCTCCATCTCATCAACCTGAACAGGAGACCGA-CCCTTGCTGCTGCTCCAGGA 852
Db 318 CCATCTCCATCTCATCAACCTGAACAGGAGACCGACCCCTTGCCCGTCTGCTCCAGGA 259
Qy 853 ATGTCAGTGTGATG 867
Db 258 ATGTCAGTGTGATG 244

RESULT 3
US-10-029-386-22242/c
; Sequence 22242, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22242
; LENGTH: 284
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR16.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: 014197, EVALUE 2.10e-01
; OTHER INFORMATION: NT HIT: g114779172, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF975008.1, EVALUE 0.00e+00
US-10-029-386-22242

Query Match 11.2%; Score 212.8; DB 16; Length 284;
Best Local Similarity 95.6%; Pred. No. 5.3e-55;
Matches 241; Conservative 0; Mismatches 7; Indels 4; Gaps 2;
Qy 617 GTCTGAACCTCCATPACTCTTTGCGATTAAACATGACAAAGGATTTGAACGGGCAGATA 676
Db 284 GTCCGAACCTCCATPACTCTTTGCGATTAAACATGACAAAGGATTTGAACGGGCAGATA 225
Qy 677 AGTTTGCTCCACCGTTTCAGACTCTTAAAGGAGTCAACGCGACAGTGAACCTCTCTTC 736
Db 224 AGTTTGCTCCACCGTTTCAGACTCTTAAAGGAGTCAACGCGACAGTGAACCTCTCTTC 168
Qy 737 TGAAGGAGTCCACGAAGGAGTGAAGCCTGTTTCAAGGAGATCAAGCCTCTCTCTGCGG 796
Db 167 TGAAGGAGTCCACGAAGGAGTGAAGCCTGTTTCAAGGAGATCAAGCCTCTCTCTGCGG 108
Qy 797 TCTCCATCTCATCAAAACCTGAACAGGAGACCGA-CCCTTGCTGCTGCTCCAGGAATG 855
Db 107 TCTCCATCTCATCAAAACCTGAACAGGAGACCGACCCCTTGCCCGTCTGCTCCAGGAATG 48
Qy 856 TCAGTGTGATG 867
Db 47 TCAGTGTGATG 36

RESULT 4
US-10-029-386-8336
; Sequence 8336, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8336
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```
; OTHER INFORMATION: MAP TO CHR16.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: SWISSPROT HIT: Q92883, EVALUE 1.20e+00
; OTHER INFORMATION: NT HIT: g115315787, EVALUE 2.00e-65
; OTHER INFORMATION: EST_HUMAN HIT: AL135642.1, EVALUE 2.00e-65
US-10-029-386-8336

Query Match      6.7%; Score 128; DB 16; Length 500;
Best Local Similarity 100.0%; Pred. No. 1.4e-28;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 237 CAGGGTCACCTGTCGTGTGCCAGTTTGAAGCCGCTCTGCAGCATGGCTTGAAGAGGAG 296
Db 341 CAGGGTCACCTGTCGTGTGCCAGTTTGAAGCCGCTCTGCAGCATGGCTTGAAGAGGAG 400

Qy 297 TCAGAGATTGGCACTACAGCGGAGCGATCAAGCAGCGAGCGGGCTTTGCCAGCAAAAC 356
Db 401 TCAGAGATTGGCACTACAGCGGAGCGATCAAGCAGCGAGCGGGCTTTGCCAGCAAAAC 460

Qy 357 CGAAACAG 364
Db 461 CGAAACAG 468

RESULT 5
US-10-029-386-22036
; Sequence 22036, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22036
; LENGTH: 124
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR16.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: NT HIT: g115315787, EVALUE 9.00e-64
; OTHER INFORMATION: EST_HUMAN HIT: AL135642.1, EVALUE 1.00e-63
; OTHER INFORMATION: SWISSPROT HIT: Q9SXU1, EVALUE 6.00e+00
US-10-029-386-22036

Query Match      6.5%; Score 124; DB 16; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.3e-27;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 GTCACTGTCTGTGTGCCAGTTTGAAGCCGCTCTGCAGCATGGCTTGAAGAGAGTCGA 300
Db 1 GTCACTGTCTGTGTGCCAGTTTGAAGCCGCTCTGCAGCATGGCTTGAAGAGAGTCGA 60

Qy 301 GGATTGGCACTACAGCGGAGCGATCAAGCAGCGAGCGGGCTTTGCCAGCAAAACCGAA 360
Db 61 GGATTGGCACTACAGCGGAGCGATCAAGCAGCGAGCGGGCTTTGCCAGCAAAACCGAA 120

Qy 361 ACAG 364
Db 121 ACAG 124

RESULT 6
US-10-276-774-948/c
; Sequence 948, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 948
; LENGTH: 3931
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-948

Query Match      3.3%; Score 63.8; DB 18; Length 3931;
Best Local Similarity 51.6%; Pred. No. 3.3e-08;
Matches 146; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

Qy 368 CCGTGTTCGTGTAAGAGGTCCTCAACAGCAGCAGCTGCAGCGCTTCTACT 427
Db 3721 CTGGCTACTGGGTGCTCGTGGTGCAATTTACTCGGAGAGAGGCCATCAAGCAGATCGAGG 3662

Qy 428 CCCTGCGCACATCGCCTCAGAGCTGGCGGGGTGGCGCTGCGCTGCGCTGCGCTCA 487
Db 3661 TCGTCAGCAGCTGGCGCACCAACCTGGGCGCAGCGCTGCTGCTGCTGCTGCTCA 3602

Qy 488 ACGAACACTCCCTGGAGCGCTACTGCATCATGTCTCTGGCCGACCGCTGCAGCTGAGCA 547
Db 3601 ACGAGAACTCCTTGGAGAGCTACTCTCGGTGTGTTCCAGAGAACCTGGGCGCTGCTGATA 3542

Qy 548 CTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCAGTATGCTTCTTACCA 607
Db 3541 AGTACTACGTCAAGAAATGCCCTGGTCTGCAGCCACGATCACCTGACCTCTTCTGACCT 3482

Qy 608 TGCAGCAGCTGTAAGCTCCTACTCTTTGGCGAATTAACATCGA 650
Db 3481 TGGTGTCCGGGCTAGAGTTCAATTCGTTTCGAGCTGGATCTGGA 3439

RESULT 7
US-09-867-701-9696/c
; Sequence 9696, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9696
; LENGTH: 263
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-9696

Query Match      3.1%; Score 59.6; DB 9; Length 263;
```

Query Match 2.9%; Score 54.6; DB 10; Length 371;
Best Local Similarity 70.2%; Pred. No. 1.8e-07;
Matches 80; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 1793 GTTCTGATGTCATTTGAGCTCCTGGATCCAGTCTTTCCCTGAAGCTGTGTTTCCTCT 1852
Db 136 GTCTTGGTGACATCTTTTGAGCCCTAGACCAAGCTTTACCTGAAGCAGAGCTACCTCA 77
QY 1853 GGACTTTTCATGATGTGAGCCCAATAAATGCTTTTCATCTTCCCTGAAAAA 1906
Db 76 GAACTTTTCAGCTATGTGAGCCCAATAAATGCTTTTCATCTTCCCTGAAAAA 23

RESULT 8

US-09-918-995-24284
; Sequence 24284, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24284
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-24284

Query Match 2.9%; Score 54.6; DB 10; Length 371;
Best Local Similarity 73.5%; Pred. No. 7.4e-06;
Matches 83; Conservative 0; Mismatches 29; Indels 1; Gaps 1;
QY 1795 TCTGTAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCCTGAAGCT-GTGTTCCTCTG 1853
Db 247 TCTGTAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCCTGAAGCT-GTGTTCCTCTG 306
QY 1854 GACTTTTCATGATGTGAGCCCAATAAATGCTTTTCATCTTCCCTGAAAAA 1906
Db 307 GACTTTTCATGATGTGAGCCCAATAAATGCTTTTCATCTTCCCTGAAAAA 359

RESULT 9

US-09-880-107-1518/c
; Sequence 1518, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1518
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA621242
US-09-880-107-1518

Query Match 2.9%; Score 54.6; DB 9; Length 446;
Best Local Similarity 73.5%; Pred. No. 8.2e-06;
Matches 83; Conservative 0; Mismatches 29; Indels 1; Gaps 1;
QY 1795 TCTGTAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCCTGAAGCT-GTGTTCCTCTG 1853
Db 113 TCTGTAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCCTGAAGCT-GTGTTCCTCTG 54
QY 1854 GACTTTTCATGATGTGAGCCCAATAAATGCTTTTCATCTTCCCTGAAAAA 1906
Db 53 GACTTTTCATGATGTGAGCCCAATAAATGCTTTTCATCTTCCCTGAAAAA 1

RESULT 10

US-09-918-995-15177
; Sequence 15177, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15177
; LENGTH: 464
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-15177

Query Match 2.8%; Score 54; DB 10; Length 464;
Best Local Similarity 57.8%; Pred. No. 1.3e-05;
Matches 96; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 365 AGCCCGTGTCTGTGTACTAGTGAAGGAGTCTCTCAACAGCAGCAGCTGCAGGGTTCT 424
Db 298 AGCCTGTCTTCTGCCCTCTCTGAAAGCTGTCAACCCCAACACATCATCTCAGAGTTGG 357
QY 425 ACTCCCTGCCCATCGCTCAGACGTGGCGGGTCCGCTGCGCTGCGTGTGCCCC 484
Db 358 AGCACCTGAGCTTTGTCAATGATGTGGCGCTGCCGGGATGGCTGGCGTGGCCCC 417
QY 485 TCAACGAACACTCCCTGGAGCGCTTACCTGCACATGTCTCTGGCGGA 530
Db 418 TGAACAATGGCTGATGAGTGCTTACCTGAAGCTGCTGCTGCAGGA 463

RESULT 11

US-09-854-867-118
; Sequence 118, Application US/09854867
; Publication No. US20030224356A1
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL H
; APPLICANT: ROGAN, PETER K
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/854,867
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 613
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 118
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)..(476)
; OTHER INFORMATION:
US-09-854-867-118

Query Match 2.8%; Score 53.2; DB 10; Length 476;
Best Local Similarity 69.5%; Pred. No. 2.3e-05;
Matches 89; Conservative 0; Mismatches 33; Indels 6; Gaps 1;

QY 1776 GAGAGAGAAATCCACAGTTCTCTGATAGTCTCATTTGAGCTCTGGATCCAGTCTTTCTT 1835
DB 313 GAGATGAGAGAGACAGATTCCTGATGACATCATTTTGGACACCTGGATCCAGCCATGCCT 372

QY 1836 GAAG-----CTGTGTTTCTCTGGACTTTTCATGTATGTAGGCCAATAAATGCTTTCA 1889
DB 373 GAAGCCAGATCTATGCTACCCCTGGACTTTTCAGTTACATGAGCCAAATAATCTCTTTT 432

QY 1890 TTCCTTGA 1897
DB 433 TTGCTTAA 440

RESULT 12
US-10-786-970A-118
; Sequence 118, Application US/10786970A
; Publication No. US2005006449A1
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/10/786, 970A
; CURRENT FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: US/09/573, 080
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 118
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: (1)..(476)
; PUBLICATION INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
; DATABASE ENTRY DATE:
; DATABASE ENTRY DATE: 1996-01-26
US-10-786-970A-118

Query Match 2.8%; Score 53.2; DB 21; Length 476;
Best Local Similarity 69.5%; Pred. No. 2.3e-05;
Matches 89; Conservative 0; Mismatches 33; Indels 6; Gaps 1;

QY 1776 GAGAGAGAAATCCACAGTTCTCTGATAGTCTCATTTGAGCTCTGGATCCAGTCTTTCTT 1835
DB 313 GAGATGAGAGAGACAGATTCCTGATGACATCATTTTGGACACCTGGATCCAGCCATGCCT 372

QY 1836 GAAG-----CTGTGTTTCTCTGGACTTTTCATGTATGTAGGCCAATAAATGCTTTCA 1889
DB 373 GAAGCCAGATCTATGCTACCCCTGGACTTTTCAGTTACATGAGCCAAATAATCTCTTTT 432

QY 1890 TTCCTTGA 1897
DB 433 TTGCTTAA 440

RESULT 13
US-10-101-510-527

; Sequence 527, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 527
; LENGTH: 7093
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-527

Query Match 2.6%; Score 50.4; DB 15; Length 7093;
Best Local Similarity 74.0%; Pred. No. 0.00065;
Matches 77; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 1795 TCCTGATAGTCTCATTTGAGCTCTCGATCCAGTCTTTCTGAGCT-GTGTTTCTCTG 1853
DB 6443 TCCTGATGACAGCGTTTGTGCCCTCGATCCAAACCGTGCCTGAAGCTAGAAATATCCCTG 6502

QY 1854 GACTTTTCATGTATGTAGCCAAATAAATGCTTTTCATCTCTTGA 1897
DB 6503 GACTTTTCAGTTATGTGAACCAATAATACCCCTTTTTCGCTTAA 6546

RESULT 14
US-10-242-355-1024
; Sequence 1024, Application US/10242355
; Publication No. US20030235831A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC003C1
; CURRENT APPLICATION NUMBER: US/10/242,355
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,897
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1024
; LENGTH: 10126
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-355-1024

Query Match 2.6%; Score 50.4; DB 17; Length 10126;
Best Local Similarity 74.0%; Pred. No. 0.00077;

us-09-155-676b-1.rnpb

Mon Jun 13 13:27:09 2005

Matches 77; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
 QY 1795 TCCTGATAGTCTCATTTGAGTCTCTGATCCAGTCTTTCTGAGCT-GTGTTCCTCTG 1853
 Db 9473 TCCTGATGACAGCGTTTGTGCCCTGATCAACCGTGCCTGAGCTAGAAATATCCCTG 9532
 QY 1854 GACTTTCATGTATGTAGCGCAATAAATGCTTTCATTCCTTGA 1897
 Db 9533 GACTTTCAGTTAAGTGAACCAATAATACCTTTTGTCTAA 9576

RESULT 15
 US-10-437-963-63143
 ; Sequence 63143, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 63143
 ; LENGTH: 550
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(550)
 ; OTHER INFORMATION: unsure at all n locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_64413C.1
 US-10-437-963-63143

Query Match 2.6%; Score 50.2; DB 19; Length 550;
 Best Local Similarity 87.3%; Pred. No. 0.00021;
 Matches 55; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 9 CACGCGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGAGGAATTCGATTCGAGG 68
 Db 21 CACCGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGAGGAATTCGAGG 80
 QY 69 CCA 71
 Db 81 GCA 83

Search completed: June 11, 2005, 06:12:51
 Job time : 726.339 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:20:34 ; Search time 4016.07 Seconds
(without alignments)
18065.048 Million cell updates/sec

Title: US-09-155-676B-1
Perfect score: 1906
Sequence: 1 cattgggtcacgcgggtggcg.....tcattctctgaaaaaaaaa 1906

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	797	41.8	1981	3	BC024293	BC024293 Homo sapi
2	672.2	35.3	730	1	AU122438	AU122438 AU122438
3	648	34.0	691	4	BM668241	BM668241 UI-E-CK1-
4	629.8	33.0	998	5	BQ064251	BQ064251 AGENCOURT
5	627.6	32.9	990	5	BQ060776	BQ060776 AGENCOURT
6	624.2	32.7	970	5	BQ065026	BQ065026 AGENCOURT
7	610.2	32.0	829	4	BG757448	BG757448 602711092
8	603.8	31.7	835	5	BQ058677	BQ058677 602712923
9	603.8	31.7	1022	5	BQ057616	BQ057616 AGENCOURT
10	600.6	31.5	843	4	BG758425	BG758425 602712623
11	600.4	31.5	892	5	BQ057008	BQ057008 AGENCOURT
12	586	30.7	609	4	BM692924	BM692924 UI-E-CK1-
13	550.8	28.9	813	4	BG684262	BG684262 602635896
14	522.6	27.4	626	5	BQ184171	BQ184171 UI-E-EJ1-
15	513.2	26.9	819	4	BG759681	BG759681 602713374
16	509.6	26.7	620	5	BQ185857	BQ185857 UI-E-EJ1-
17	489	25.7	769	4	BG684980	BG684980 602636729
18	482.8	25.3	504	4	BM697878	BM697878 UI-E-DX0-
19	482	25.3	914	4	BG340557	BG340557 602462123
20	481.6	25.3	782	4	BG759163	BG759163 602710820
21	475.4	24.9	951	4	BG685372	BG685372 602637216
22	474.4	24.9	744	9	AY399018	AY399018 Homo sapi
23	467.8	24.5	1064	5	BQ065934	BQ065934 AGENCOURT
24	461	24.2	501	1	AL135642	AL135642 DKF2p762K

25	454.8	23.9	737	4	BF975008	BF975008 602244582
26	454.2	23.8	1075	5	BQ058038	BQ058038 AGENCOURT
c 27	453	23.8	855	5	BX760749	BX760749
28	447	23.5	1051	5	BQ061744	BQ061744 AGENCOURT
29	437.2	22.9	576	2	AW974284	AW974284 EST386387
30	436	22.9	965	4	BG757633	BG757633 602711219
31	431.4	22.6	822	4	BM478075	BM478075 AGENCOURT
c 32	427	22.4	439	1	AI369689	AI369689 qy71908.x
33	410.8	21.6	682	9	AY399019	AY399019 Pan trogl
34	385.4	20.2	744	9	AY399020	AY399020 Mus muscu
35	382.4	20.1	419	1	AA593354	AA593354 nn06h08.g
36	376.8	19.8	850	5	BQ642454	BQ642454 AGENCOURT
37	372.8	19.6	430	2	AW03206	AW03206 UI-HF-BKO
38	372.4	19.5	522	2	AW044949	AW044949 um15h06.y
39	370	19.4	506	5	BX283539	BX283539
40	358.8	18.8	520	2	AW503340	AW503340 UI-HF-BNO
41	345	18.1	847	7	CR579211	CR579211
42	343	18.0	478	5	BX283759	BX283759
43	341.6	17.9	646	1	AL965834	AL965834
44	323.2	17.0	602	1	AL959249	AL959249
45	317.6	16.7	498	4	BG759800	BG759800 602711706

ALIGNMENTS

RESULT 1
BC024293
LOCUS BC024293 1981 bp mRNA linear HTC 04-MAR-2003
DEFINITION Homo sapiens, clone IMAGE:5088007, mRNA.
ACCESSION BC024293
VERSION BC024293.1 GI:22028130
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1981)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (26-PEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Heiao, Martin Krzywnski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 42 Row: m Column: 22
This clone has the following problem: retained intron.

FEATURES

source
1. .1981
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5088007"
/tissue_type="Lymph, lymphoma"

DEFINITION AUI22438 MAMMA1 Homo sapiens cDNA clone MAMMA1002380 5', mRNA sequence.
ACCESSION AUI22438
VERSION AUI22438.1 GI:10937708
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 730)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
FEATURES
source Location/Qualifiers
1..730
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MAMMA1002380"
/tissue_type="mammary gland"
/clone_lib="MAMMA1"
/note="Vector: pME18SFL3"
ORIGIN
Query Match 35.3%; Score 672.2; DB 1; Length 730;
Best Local Similarity 97.9%; Pred. No. 8.5e-171; Indels 2; Gaps 2;
Matches 696; Conservative 0; Mismatches 13;
QY 77 GCCGGCGCGCGCGCANGCACCGGCCCGGGGAGNCCATGAGCGGATCNCNGAACN 136
Db 22 GCGCGCGCGCGCGCGCAGGACCGGCCCGGGGAGGAGCCATGAGCGGATCAGAAACA 81
QY 137 ATGACAAAGACAATTTCTGCTGGAGCGACTGCTGGATGAGTGAACAGTGCAGATCC 196
Db 82 ATGACAAAGACAATTTCTGCTGGAGCGACTGCTGGATGAGTGAACAGTGCAGATCC 141
QY 197 GCTTTNGAGGAGAAAGAGAGATTCCTCGGATTCGACAGAGGATTCAGGATTCGACAC 256
Db 142 GCTTTGGAGGAGAAAGAGAGATTCCTCGGATTCGACAGAGGATTCAGGATTCGACAC 201
QY 257 CCCAGTTTGAAGCGCTCTCGACGATGCTTGAAGAGAGTTCGAGGATTCGACACACAG 316
Db 202 CCCAGTTTGAAGCGCTCTCGACGATGCTTGAAGAGAGTTCGAGGATTCGACACACAG 261
QY 317 GCGCAGCGATCAAGCAGCAGCGGGCTTTCGACGAAACCGAAACAGAGCCCGTGTCT 376
Db 262 GCGCAGCGATCAAGCAGCAGCGGGCTTTCGACGAAACCGAAACAGAGCCCGTGTCT 321
QY 377 GGTACTACGTGAAGAGGTTCCTCAACAGCAGCAGCTCAGCGCTTCTACTCCCTGCGCC 436
Db 322 GGTACTACGTGAAGAGGTTCCTCAACAGCAGCAGCTCAGCGCTTCTACTCCCTGCGCC 381
QY 437 ACATCGCTCAGACGTGGGCGGGGTTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 496
Db 382 ACATCGCTCAGACGTGGGCGGGGTTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 441
QY 497 CCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCGAGGCTGAGCACTTTTATG 556
Db 442 CCCTGGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCGAGGCTGAGCACTTTTATG 501

ORIGIN
Query Match 41.8%; Score 797; DB 3; Length 1981;
Best Local Similarity 97.0%; Pred. No. 1.5e-204;
Matches 855; Conservative 0; Mismatches 20; Indels 6; Gaps 4;
QY 362 CAGAGCCCGTCTCTGGTACTACGTGAAGGAGTCCCTCAACAGCAGCAGCTGC-AGCGC 420
Db 362 CAGAGCCCGTCTCTGGTACTACGTGAAGGAGTCCCTCAACAGCAGCAGCTGCAGAGCGC 421
QY 421 TTCTACTCCCTCGCCACATCGCTCAGACGTGGCGGGTCCGGCTCGCTGCGCTGT 480
Db 422 TTCTACTCCCTCGCCACATCGCTCAGACGTGGCGGGTCCGGCTCGCTGCGCTGT 481
QY 481 GCCCTCAACGAACAATCCCTGAGCGCTACCTGCACATGCTCCTGGCGGCGCGCTGCAGG 540
Db 482 GCCCTCAACGAACAATCCCTGAGCGCTACCTGCACATGCTCCTGGCGGCGCGCTGCAGG 541
QY 541 CTGAGCACTTTTATGAAGACTGTCTTTTGTGATGATGAAGAAAGTCCAGTATGCTT 600
Db 542 CTGAGCACTTTTATGAAGACTGTCTTTTGTGATGATGAAGAGGTCCAGTATGCTT 601
QY 601 CCTACCATGGCAGCGGTCTGAATCCATACCTTTTGGCATTAACATCGACAACAGAT 660
Db 602 CCTACCATGGCAGCGGTCCGAATCCATACCTTTTGGCATTAACATCGACAACAGAT 661
QY 661 TTGAACGGCAGAGTAAGTTTCTCCACCGTTTCAGACCTCTTAAAGAGTCAACGCAG 720
Db 662 TTGAACGGCAGAGTAAGTTTCTCCACCGTTTCAGACCTCTTAAAGAGTCAACGCAG 721
QY 721 AAGCTCACCTCTCTGCTGAAGAGTCCACGCAAGAGTGAGCAGCTGTTCAGGGAGATC 780
Db 722 AAGTGA---CCTTCTGAAGAGTCCACGCAAGAGTGAGCAGCTGTTCAGGGAGATC 778
QY 781 ACAGCTCTCTGCTGCTCTCCATCTCATCAACCTGACAGGAGCAGCA-CCCTTGCT 839
Db 779 ACAGCTCTCTGCTGCTCTCCATCTCATCAACCTGACAGGAGCAGCAACCCCTTGCT 838
QY 840 GTCTGTCCAGGAATGTCACTGCTGATGCAATGCAAAAGAGGAGCGGAAGAAAGAAAG 899
Db 839 GTCTGTCCAGGAATGTCACTGCTGATGCCAATGCAAAAGAGGAGCGGAAGAAAGAAAG 898
QY 900 AAGTGAACCAATAATCTCATTTGATGAGGAAGATGAGCAGAACTCTGGGACGTG 959
Db 899 CAAGTGACCAACATTAATCTCATTTGATGAGGAAGATGAGCAGAACTCTGGGACATG 958
QY 960 TTTTAAAGACACTGGGCGAGGAGCTCAGAGCAAACTCCGACCGCTCTCTGTC 1019
Db 959 TTTTAAAGACACTGGGCGAGGAGCTCAGAGCAAACTCCGACCACTCTCTGTC 1018
QY 1020 AATATCATGTCGCTTTTGAAGCCCTTCGGGCTTAATCTCAATGGAA-TCAGACAGC 1078
Db 1019 AATATCATGTCGCTTTTGAAGCCCTTCGGGCTTAATCTCAATGGAACTCAGACAGC 1078
QY 1079 AACTCATGGAATAATGATTCCTGTCTTTGAAACGGGAGTTTGGTACCAAGAGCTTGAT 1138
Db 1079 AACTCATGGAATAATGATTCCTGTCTTTGAAACGGGAGTTTGGTACCAAGAGCTTGAT 1138
QY 1139 GTGAAAGCATCGATGATGAGATGTGGATGAAACCAAGATACGCTGTATGAAACTCA 1198
Db 1139 GTGAAAGCATCGATGATGAGATGTGGATGAAACCAAGATACGCTGTATGAAACTCA 1198
QY 1199 TCAGGAAGAAAGCAGAGGGCCACTCGGAGTCGCCCGAGAA 1239
Db 1199 TCAGGAAGAAAGCAGAGGGCCACTCAGAGTCGCCCGAGAA 1239

QY 557 AAGACTGGTCTTTTGTGATGATGAAGAAAGGTCAGTATGCTTCCTACCATGCGACGAG 616
 Db 502 AAGACTGGTCTTTTGTGATGATGAAGAAAGGTCAGTATGCTTCCTACCATGCGACGAG 561
 QY 617 GTCTGAAGTCTCATACTCTTTTCGATTAACTGCAACAGGATTTGAACGGGCGAGGTA 676
 Db 562 GTCTGAAGTCTCATACTCTTTTCGATTAACTGCAACAGGATTTGAACGGGCGAGGTA 621
 QY 677 AGTTTGTCTCCACCGTTTCAGACCTCTTTAAAGGATCAACGCGAGAACTGACCTTCCTTGC 736
 Db 622 AGTTTGTCTCCACCGTTTCAGACCTCTTTAAAGGATCAACGCGAGAACTGACCTTCCTTGC 680
 QY 737 TGAAGGAGTCCACCAAGGAGTGAAGCCCTGTTTCAGGAGATCACAGCCT 787
 Db 681 TGAAGGAGTCCACCAAGGAGTGANCA-CCTGTTCAAGGAGATCACAGCCT 730

RESULT 3

BM668241/c

LOCUS

DEFINITION

UI-E-CK1-afk-h-18-0-UI.s2 UI-E-CK1 Homo sapiens cDNA clone

UI-E-CK1-afk-h-18-0-UI 3', mRNA sequence.

BM668241

BM668241.1 GI:18976072

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 691)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

8889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Forward

POLYA=Yes.

Location/Qualifiers

1..691

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-E-CK1-afk-h-18-0-UI"

/tissue_type="Retina Foveal and Macular"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-E-CK1"

/notes="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-E-CK1 is a normalized cDNA library containing the

following tissue(s): Retina Foveal and Macular. The

library was constructed according to Bonaldo, Lennon and

Soares, Genome Research, 6:791-806, 1996. First strand

cDNA synthesis was primed with an oligo-dT primer

containing a Not I site. Double stranded cDNA was ligated

to an EcoR I adaptor, digested with Not I, and cloned

directionally into pT7T3-Pac vector. The oligonucleotide

used to prime the synthesis of first-strand cDNA contains

a library tag sequence that is located between the Not I

site and the (dT)18 tail. The sequence tag for this library is GTCC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
 TAG TISSUE=Foveal and Macular Retina
 TAG_LIB=UI-E-CK1
 TAG_SEQ=GTCC"

ORIGIN

Query Match 34.0%; Score 648; DB 4; Length 691;
 Best Local Similarity 99.6%; Pred. No. 3.1e-164;
 Matches 681; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
 QY 1225 GAGTGTGCCCGAGAGCCACTGTGAAGGAAACACTGCTCTTCCAGATGACACAGTGGGC 1284
 Db 691 GAGTGTGCCCGAGAGCCACTGTGAAGGAAACACTGCTCTTCCAGATGACACAGTGGGC 632
 QY 1285 TCCGCTGAAAGTGTCTGCACATGACTCCGACATCTCTTCCCTGTCACTGCGGTGGGCTC 1344
 Db 631 TCCGCTGAAAGTGTCTGCACATGACTCCGACATCTCTTCCCTGTCACTGCGGTGGGCTC 572
 QY 1345 CTACAGCCCGAGAGATGCCCTTCCGAAAGCTTGGAGAACGGACAGACAGGAGACCA 1404
 Db 571 CTACAGCCCGAGAGATGCCCTTCCGAAAGCTTGGAGAACGGACAGACAGGAGACCA 512
 QY 1405 CGTTCTCCGGATCCTGGACTTCGGTACAGTGTGGAGCCAGCTCTCCAGGCCACGGAAG 1464
 Db 511 CGTTCTCCGGATCCTGGACTTCGGTACAGTGTGGAGCCAGCTCTCCAGGCCACGGAAG 452
 QY 1465 TCCTCTGAGGAGCTGTGTTA-CCTTCTGCTCAGTGTCCAGAGTCCATGACAAATTAGTGAAT 1523
 Db 451 TCCTCTGAGGAGCTGTGTTA-CCTTCTGCTCAGTGTCCAGAGTCCATGACAAATTAGTGAAT 392
 QY 1524 GCGCCAGGCGCTGTGGCCATGATGAACAGAGGAGTGTGGAGGAGGAGAACAGATC 1583
 Db 391 GCGCCAGGCGCTGTGGCCATGATGAACAGAGGAGTGTGGAGGAGGAGAACAGATC 332
 QY 1584 ACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGCTCCGCGAAGAGGTGGA 1643
 Db 331 ACTGCGAAACCTGCTCGACGGTGAGATGGAGCACTCAGCCGCTCCGCGAAGAGGTGGA 272
 QY 1644 CACCTTGAAGGAGGAGTGTGCTGAAACAGAGGAGGCGGAGGATGAGGTCCAGGCGCT 1703
 Db 271 CACCTTGAAGGAGGAGTGTGCTGAAACAGAGGAGGCGGAGGATGAGGTCCAGGCGCT 212
 QY 1704 GGCAGCTATCTTTGCTATTTTGTGAGGAGATTTAAACCCAGTGTGAGAACCATGTGCTG 1763
 Db 211 GGCAGCTATCTTTGCTATTTTGTGAGGAGATTTAAACCCAGTGTGAGAACCATGTGCTG 152
 QY 1764 GAGAAATGGAGGAGAGAGAAATCCAAACAGTTCCTGTATGCTCTCATTTGAGCTCCTGGAT 1823
 Db 151 GAGAAATGGAGGAGAGAGAAAT-CAACAGTTCCTGTATGCTCTCATTTGAGCTCCTGGAT 93
 QY 1824 CAGTCTTTCCTGAGAGCTGTGTTTCTCTGAGCTTTTTCATGTATGTGAG-CCAATAAATT 1882
 Db 92 CAGTCTTTCCTGAGAGCTGTGTTTCTCTGAGCTTTTTCATGTATGTGAGCCAATAAATT 33
 QY 1883 GCTTTTCATCTTGTGAAAAA 1906
 Db 32 GCTTTTCATCTTGTGAAAAA 9

RESULT 4

BM664251

LOCUS

DEFINITION

AGENCOURT_6854182 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5925960

5', mRNA sequence.

BM664251

BM664251.1 GI:19892719

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 998)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2098 row: i column: 01
High quality sequence stop: 692.
Location/Qualifiers

FEATURES
source

1..998
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5925960"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH-MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH-MGC
Library."

ORIGIN

Query Match 33.0%; Score 629.8; DB 5; Length 998;
Best Local Similarity 96.7%; Pred. No. 3e-159;
Matches 675; Conservative 0; Mismatches 18; Indels 5; Gaps 3;

QY 544 AGCAGCTTTTATGAGACTGGTCTTTTGTGATGATGAAGAAGTCCAGTATGCTTCT 603
DB 92 AGTACTTTTATGAGACTGGTCTTTTGTGATGATGAAGAAGTCCAGTATGCTTCT 151
QY 604 ACCATGGCAGAGTCTGAATCCATCTCTTTGGATTAACATCGACACAGGATTTG 663
DB 152 ACCATGGCAGAGTCCGAATCCATCTCTTTGGATTAACATCGACACAGGATTTG 211
QY 664 AACGGCAGAGTAAGTTTCTCCACCGTTTCAGACTCTTAAAGAGTCAACGAGAAC 723
DB 212 AACGGCAGAGTAAGTTTCTCCACCGTTTCAGACTCTTAAAGAGTCAACGAGAAC 271
QY 724 GTGACCTCTTGTGAAGAGTCCACGCAAGAGTGAAGAGTCTGTTTCAAGGAGATCACA 783
DB 272 GTGA---CTTGTGAAGAGTCCACGCAAGAGTGAAGAGTCTGTTTCAAGGAGATCACA 328
QY 784 GCCTCTCTGCTCTCCATCTTCATCAACCTGAACGAGGACCGA-CCCTTGGCTGTC 842
DB 329 GCCTCTCTGCCATCTCCATCTTCATCAACCTGAACGAGGACCGA-CCCTTGGCTGTC 388
QY 843 GTGTCCAGGAATGTGAGTGTGATGCAATGCAAAAGAGGAGCGGAAGAAAGAA 902
DB 389 GTGTCCAGGAATGTGAGTGTGATGCAATGCAAAAGAGGAGCGGAAGAAAGAA 448
QY 903 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGACGTTT 962
DB 449 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGACGTTT 508
QY 963 AAAAGACACCTGGGCGAGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAT 1022
DB 509 AAAAGACACCTGGGCGAGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAT 568
QY 1023 ATCATGTCCGCCTTTTGAAGCCCTTTCGGGCTTAACCTCAATGAA-TCAGAGCAAC 1081

Db 569 ATCATGTCCGCTTTGAAAGCCCTTCGGGCCAAACTCCATGGAAGTCAGAGCAGCAAC 628
QY 1082 TCATGAAATTAATGATTCCTCTTTGAAAGGGAGTTTGGTACAGAGCTTGTATG 1141
Db 629 TCTGGAATTAATGATTCCTCTTTGAAAGGGAGTTTGGTACAGAGCTTGTATG 688
QY 1142 AAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1201
Db 689 AAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 748
QY 1202 GGAAGCAAGCAGAGGCGCACTCGGAGTCCCGAGAA 1239
Db 749 GGACGAGCAGACAGGCGCACTCAGAGTCCCGAGAA 786

RESULT 5
BQ060776 990 bp mRNA linear EST 02-APR-2002
LOCUS AGENCOURT 7049038 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5816452
DEFINITION 5', mRNA sequence.
ACCESSION BQ060776
VERSION BQ060776.1 GI:19884191
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaaps-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2073 row: f column: 05
High quality sequence stop: 686.
Location/Qualifiers

FEATURES
source

1..990
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5816452"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH-MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH-MGC
Library."

ORIGIN

Query Match 32.9%; Score 627.6; DB 5; Length 990;
Best Local Similarity 96.6%; Pred. No. 1.2e-158;
Matches 674; Conservative 0; Mismatches 19; Indels 5; Gaps 3;

QY 544 AGCAGCTTTTATGAGACTGGTCTTTTGTGATGATGAAGAAGTCCAGTATGCTTCT 603
DB 98 AGTACTTTTATGAGACTGGTCTTTTGTGATGATGAAGAAGTCCAGTATGCTTCT 157
QY 604 ACCATGGCAGAGTCTGAATCCATCTCTTTCGATTAACATCGACACAGGATTTG 663

```
Db 158 ACCATGGCAGCAGGTCCGAACCTCCATCTCTTTGGATTAACTTGAACAACAGGATTTG 217
Qy 664 AACGGGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGAGAGTCAACGCAAC 723
Db 218 AACGGGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGAGAGTCAACGCAAT 277
Qy 724 GTGACCTCTCTTGTCTGAAGGAGTCCACGACAGGAGTCCAGGAGTCCAGGAGATCACA 783
Db 278 GTGA---CCTTGTCTGAAGGAGTCCACGACAGGAGTCCAGGAGTCCAGGAGATCACA 334
Qy 784 GCCTCTCTGCGCTCTCCATCTCTCAACCTCAACAGGAGACCGA-CCTTTGCTCTGTC 842
Db 335 GCCTCTCTGCGCTCTCCATCTCTCAACCTCAACAGGAGACCGA-CCTTTGCTCTGTC 394
Qy 843 GTGTCCAGGAATGTCAAGTCTGATGCTCAATGCAAAAGAGGCGGAAGAAAGAA 902
Db 395 GTGTCCAGGAATGTCAAGTCTGATGCTCAATGCAAAAGAGGCGGAAGAAAGAA 454
Qy 903 GTGACCAACATATCTCATTTGATGAGGAGATGAGCAGACCTCTGCGGAGCTTTT 962
Db 455 GTGACCAACATATCTCATTTGATGAGGAGATGAGCAGACCTCTGCGGAGCTTTT 514
Qy 963 AAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT 1022
Db 515 AAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT 574
Qy 1023 ATCATGTCCGCTTTGAAAGCCCTTCGGGCTTAACCTCAATGGAATCGAAGCAGCAAC 1081
Db 575 ATCATGTCCGCTTTGAAAGCCCTTCGGGCTTAACCTCAATGGAATCGAAGCAGCAAC 634
Qy 1082 TCATGGAATATGATTCCTCTCTTTGAAAGCCCTTCGGGCTTAACCTCAATGGAATCGAAGCAGCAAC 1141
Db 635 TCATGGAATATGATTCCTCTCTTTGAAAGCCCTTCGGGCTTAACCTCAATGGAATCGAAGCAGCAAC 694
Qy 1142 AAAAGCATCATGATGAAGATGTGATGGAAGCAACGAGATGATGATGCAAACTCATCA 1201
Db 695 AAAAGCATCATGATGAAGATGTGATGGAAGCAACGAGATGATGATGCAAACTCATCA 754
Qy 1202 GGAAGGAAGCAGAGGGGCCACTCGGAGTCGCGCCGAGAA 1239
Db 755 AGACGGAAGCAGAGGGGCCACTCAGAGTCGCGCCGAGAA 792
```

```
RESULT 6
BQ065026
LOCUS
DEFINITION
AGENCOURT_685627 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5929505
5', mRNA sequence.
ACCESSION
BQ065026
VERSION
BQ065026.1 GI:19894072
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 970)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LCM2107 row: 1 column: 18
High quality sequence stop: 676.
Location/Qualifiers
1..970
/organism="Homo sapiens"
```

FEATURES
source

RESULT 7
BG757448
LOCUS

BG757448 829 bp mRNA linear EST 15-MAY-2001

```
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="IMAGE:5929505"  
/tissue_type="lymphoma, cell line"  
/lab_host="DH10B (phage-resistant)"  
/clone_lib="NIH_MGC_99"  
/note="Organ: lymph. Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH_MGC  
Library."
```

ORIGIN

```
Query Match 32.7%; Score 624.2; DB 5; Length 970;  
Best Local Similarity 96.7%; Pred. No. 9.9e-158;  
Matches 670; Conservative 0; Mismatches 18; Indels 5; Gaps 3;  
Qy 544 AGCAGCTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCCAGTATGCTTCCT 603  
Db 92 AGTACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGGTCCAGTATGCTTCCT 151  
Qy 604 ACCATGGCAGCAGGTCTGAACCTCACTCTTTTGGATTAACTCGAACAAGAGATTG 663  
Db 152 ACCATGGCAGCAGGTCCGAACCTCACTCTTTTGGATTAACTCGAACAAGAGATTG 211  
Qy 664 AACGGGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGAGTCAACGCAAC 723  
Db 212 AACGGGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGAGTCAACGCAAT 271  
Qy 724 GTGACCTCTCTGCTGAAGGAGTCCACGCAAGGAGTGAAGAGTCTGTTTCAGGAGATCACA 783  
Db 272 GTGA---CCTTGTCTGAAGGAGTCCACGCAAGGAGTGAAGAGTCTGTTTCAGGAGATCACA 328  
Qy 784 GCCTCTCTGCGCTCTCTCATCTCTCATCAACCTGAACAGGAGACCGA-CCTTTGCTCTGTC 842  
Db 329 GCCTCTCTGCGCTCTCTCATCTCTCATCAACCTGAACAGGAGACCGA-CCTTTGCTCTGTC 388  
Qy 843 GTGTCCAGGAATGTCAAGTCTGATGCTCAATGCAAAAGAGGCGGAAGAAAGAA 902  
Db 389 GTGTCCAGGAATGTCAAGTCTGATGCTCAATGCAAAAGAGGCGGAAGAAAGAA 448  
Qy 903 GTGACCAACATATCTCTCATCTCTCATCAACCTGAACAGGAGACCGA-CCTTTGCTCTGTC 962  
Db 449 GTGACCAACATATCTCTCATCTCTCATCAACCTGAACAGGAGTGAAGAGTCTTGGGGACATGTTT 508  
Qy 963 AAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT 1022  
Db 509 AAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT 568  
Qy 1023 ATCATGTCCGCTTTGAAAGCCCTTCGGGCTTAACCTCAATGGAATCGAAGCAGCAAC 1081  
Db 569 ATCATGTCCGCTTTGAAAGCCCTTCGGGCTTAACCTCAATGGAATCGAAGCAGCAAC 628  
Qy 1082 TCATGGAATATGATTCCTCTCTTTGAAAGCCCTTCGGGCTTAACCTCAATGGAATCGAAGCAGCAAC 1141  
Db 629 TCATGGAATATGATTCCTCTCTTTGAAAGCCCTTCGGGCTTAACCTCAATGGAATCGAAGCAGCAAC 688  
Qy 1142 AAAAGCATCATGATGAAGATGTGATGGAAGCAACGAGATGATGATGCAAACTCATCA 1201  
Db 689 AAAAGCATCATGATGAAGATGTGATGGAAGCAACGAGATGATGATGCAAACTCATCA 748  
Qy 1202 GGAAGGAAGCAGAGGGGCCACTCGGAGTCGCGCC 1234  
Db 749 GGAAGGAAGCAGAGGGGCCACTCAAGTCGCCC 781
```

DEFINITION 602711092F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4851502 5',
mRNA sequence.
ACCESSION BG757448
VERSION BG757448.1 GI:14068101
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 829)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI894 row: 9 column: 23
High quality sequence stop: 801.
Location/Qualifiers
1. 829
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4851502"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site:1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 32.0%; Score 610.2; DB 4; Length 829;
Best Local Similarity 95.9%; Pred. No. 6a-154;
Matches 670; Conservative 0; Mismatches 23; Indels 6; Gaps 4;

QY 544 AGCACTTTTATGAAGACTGGTCTTTTGTGATGAAGAAAGTCCAGTATGCTTCCT 603
DB 133 AGTACTTTTATGAAGACTGGTCTTTTGTGATGAAGAGGTCCAGTATGCTTCCT 192

QY 604 ACCATGGCAGGAGTCTGAATCCATCTCTTTGGGATTAAATCATCGACAAACAAGATTG 663
DB 193 ACCATGGCAGGAGTCCGAATCCATCTCTTTGGGATTAAATCATCGACAAACAAGATTG 252

QY 664 AACGGGCAGAGTAAGTTTGTCTCCACCTTTTCAGACCTTTAAAGGAGTCAACGAGAAC 723
DB 253 AACGGGCAGAGTAAGTTTGTCTCCACCTTTTCAGACCTTTAAAGGAGTCAACGAGAAAT 312

QY 724 GTGACCTCTCTGCTGAAGAGTCCACGAAGAGTGAAGAGTCTTTTCAGGAGATCACA 783
DB 313 GTGA---CTTTGCTGAAGAGTCCACGACGAGAGTGAAGAGTCTTTTCAGGAGATCACA 369

QY 784 GCCTCCTCTGCGGTCTCATCTCTATCAAACTGAAACAGGAGACCGA-CCCTTGCCTGTC 842
DB 370 GCCTCCTCTGCGATCTCCATCTCATCAAACTGAAACAGGAGACCGACCCCTTGCCTGTC 429

QY 843 GTGTCAGGAATGTCTGAGTCCCAAAATGCAAAAGAGGCGGAGCAAGAAAGAAA 902
DB 430 GTGTCAGGAATGTCTGAGTCCCAAAATGCAAAAGAGGCGGAGCAAGAAAGAAA 489

QY 903 GTGACCAACATAATCTCTATTTTGTATGAGGAAGATGACAGAACTCTTGGGGACGTGTTT 962

DB 490 GTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTT 549

QY 963 AAAAAGACACTCTGGGCGAGGGAGAGCTCAGAGACAACTCCGACCGCTCCTCTGTCAAT 1022

DB 550 AAAAAGACACTCTGGGCGAGGGAGAGCTCAGAGACAACTCCGACCACTCTCTCTGTCAAT 609

QY 1023 ATCATGTCCGCTTTTGAAGCCCTTCGGGCTTAACCTCAATGGAA-TCAGAGCAGCAAC 1081

DB 610 ATCATGTCCGCTTTTGAAGCCCTTCGGGCTTAACCTCAATGGAACTCAGAGCAGCAAC 669

QY 1082 TCATGGAAAATTGATTCCTCTGTTTGAACGGGAGGTTTGGGTACCAAGACCTTGTATGTG 1141

DB 670 TCGTGGAAAATTGATTCCTCTGTTTGAACAGGAGGTTTGGGTACCAAGACCTTGTATGTG 729

QY 1142 AAAAGCATCGATGATGAAGATGTGGATGAAACCAAGATGACGTGTATGG-AAACTCATC 1200

DB 730 AAAAGCATCGATGATGAAGATGTGGATGAAACCAAGATGACGTGTATGGAAAACCTCATC 789

QY 1201 AGGAAGGAAGCACAGGGGCCACTCGAGTCGCCCGAGAA 1239

DB 790 AGGACGGAAGCACATGTGGCACTCAGAGTCGGGCGAGAA 828

RESULT 8
BG758677 835 bp mRNA linear EST 15-MAY-2001
LOCUS 602712923F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853188 5',
mRNA sequence.
ACCESSION BG758677
VERSION BG758677.1 GI:14069330
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 835)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI698 row: n column: 05
High quality sequence stop: 819.
Location/Qualifiers
1. 835
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4853188"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site:1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

FEATURES
source

ORIGIN
Query Match 31.7%; Score 603.8; DB 4; Length 835;
Best Local Similarity 96.3%; Pred. No. 3.3e-152;

Matches 673; Conservative 0; Mismatches 17; Indels 9; Gaps 5;

544 AGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCAGATGCTTCCT 603
|||
Db |||

31 AGTACTTTTATGAAGACTGGTCTTTGTGATGATGAAGAGAGGTCAGATGCTTCCT 90
|||
Qy |||

604 ACCATGGCAGCAGGTCTGAACCTCATCTCTTTGCGATTAACATCAGCAACAGGATTTG 663
|||
Db |||

91 ACCATGGCAGCAGGTCCGAATCCATCTCTTTGCGATT---ATTGACAACAGGATTTG 147
|||
Qy |||

664 AACGGGCGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAGAGAGTCAACGCGAGAAC 723
|||
Db |||

148 AACGGGCGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAGAGAGTCAACGCGAAT 207
|||
Qy |||

724 GTGACCTCTTGTGTAAGGAGTCCACGCAAGAGTGAAGAGTCTTTTTCAGGGAGATCACA 783
|||
Db |||

208 GTGA---CCTTGTCTGAAGAGTCCACGCAAGAGTGAAGAGTCTTTTCAGGGAGATCACA 264
|||
Qy |||

784 GCTCTCTCTGCGTCTCCATCTCTCATCAACCTGAACCTGAACCTGAACCTGAACCTGAAC 842
|||
Db |||

265 GCCTCTCTGCGATCTCCATCTCTCATCAACCTGAACCTGAACCTGAACCTGAACCTGAAC 324
|||
Qy |||

843 GTCTCAGGAATCTCAGTCTGATGCCAATGCAAAAGGAGCGGAAGAAAGAAAGAA 902
|||
Db |||

325 GTCTCAGGAATCTCAGTCTGATGCCAATGCAAAAGGAGCGGAAGAAAGAAAGCA 384
|||
Qy |||

903 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTCGGACAGTGT 962
|||
Db |||

385 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTCGGACAGTGT 444
|||
Qy |||

963 AAAAAAGACCTGGGAGGAGAGTCAAGGACAACTCCGACCGCTCTCTGTCAAT 1022
|||
Db |||

445 AAAAAAGACCTGGGAGGAGAGTCAAGGACAACTCCGACCGCTCTCTGTCAAT 504
|||
Qy |||

1023 ATCATCTCGCCTTTGAAAGCCCTTCGGGCTTAACCTCAAGTGAAGTCAAGAGCAGCAAC 1081
|||
Db |||

505 ATCATCTCGCCTTTGAAAGCCCTTCGGGCTTAACCTCAAGTGAAGTCAAGAGCAGCAAC 564
|||
Qy |||

1082 TCATGAAATTTGATTCCTCTCTTTGAAAGCCCTTCGGGCTTAACCTCAAGTGAAGTCAAGAG 1141
|||
Db |||

565 TCGTGAATTTGATTCCTCTCTTTGAAAGCCCTTCGGGCTTAACCTCAAGTGAAGTCAAGAG 624
|||
Qy |||

1142 -AAAGCATCGATGATGAAGATGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1200
|||
Db |||

625 AAAAAAGCATGATGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 684
|||
Qy |||

1201 AGAAGGAGCAGCAGGCGCACTCGAGTCTGCGCCGAGAA 1239
|||
Db |||

685 AGAAGGAGCAGCAGGCGCACTCGAGTCTGCGCCGAGAA 723
|||

RESULT 9
BQ057616
LOCUS
DEFINITION BQ057616 1022 bp mRNA linear EST 29-MAR-2002
5', mRNA sequence.
ACCESSION BQ057616
VERSION BQ057616.1 GI:19816956
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1022)
NIH-MGC <http://mgi.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLCM2065 row: a column: 24
High quality sequence stop: 641.
Location/Qualifiers
1..1022
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5813279"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 31.7%; Score 603.8; DB 5; Length 1022;
Best Local Similarity 96.0%; Pred. No. 3.5e-152;
Matches .652; Conservative 0; Mismatches 22; Indels 5; Gaps 3;

544 AGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGTCCAGTATGCTTCCT 603
Db 70 AGTACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGATTTCCAGTATGCTTCCT 129
Qy 604 ACCATGGCAGCAGGTCTGAACCTCATCTCTTTGCGATTAACATCAGCAACAGGATTTG 663
Db 130 ACCATGGCAGCAGGTCCGAATCCATCTCTTTGCGATTAACTGACAAAGGATTTG 189
Qy 664 AACGGGCGAGTAAGTTTGTCTCCACCGTTTCAGACCTCTTAAGAGAGTCAACGCGAGAAC 723
Db 190 AACGGGCGAGTAAGTTTGTCTCCACCGTTTCAGACCTCTTAAGAGAGTCAACGCGAAT 249
Qy 724 GTGACCTCTTCTGTAAGAGTCCACGCAAGAGTGAAGAGTCTTTTCAGGGAGATCACA 783
Db 250 GTGA---CCTTGTCTGAAGAGTCCACGCAAGAGTGAAGAGTCTTTTCAGGGAGATCACA 306
Qy 784 GCCTCTCTGCGTCTCCATCTCTCAACCTGAACCTGAACCTGAACCTGAACCTGAACCTGAAC 842
Db 307 GCCTCTCTGCGTCTCCATCTCTCAACCTGAACCTGAACCTGAACCTGAACCTGAACCTGAAC 866
Qy 843 GTGTCAGGAATGTGATGCTGATGCCAATGCAAAAGGAGCGGAAGAAAGAAAGAA 902
Db 367 GTGTCAGGAATGTGATGCTGATGCCAATGCAAAAGGAGCGGAAGAAAGAAAGCA 426
Qy 903 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTCGGAGCAGTGT 962
Db 427 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTCGGAGCAGTGT 486
Qy 963 AAAAAAGACCTGGGCGAGGAGAGTCTGAGAGGCACTCCGACCGCTCTCTGTCAAT 1022
Db 487 AAAAAAGACCTGGGCGAGGAGAGTCTGAGAGGCACTCCGACCGCTCTCTGTCAAT 546
Qy 1023 ATCATGTCCGCTTTGAAAGCCCTTCGGGCTTAACCTCAATGGA-TCAGAGCAGCAAC 1081
Db 547 ATCATGTCCGCTTTGAAAGCCCTTCGGGCTTAACCTCAATGGAAGTTCAGAGCAGCAAC 606
Qy 1082 TCATGAAATTTGATTCCTCTCTTTGAAAGCCGAGGATTTGGGTACAGAACTGTGATGTG 1141
Db 607 TCGTGAATTTGATTCCTCTCTTTGAAAGCCGAGGATTTGGGAGCCAGAACTGTGATGTG 666
Qy 1142 AAAAAAGCATGATGAAGATGTGATGAAGAGAGATGAGCTGTATGGAACATCA 1201
Db 667 AAAAAAGCATGATGAAGATGTGATGAAGAGAGATGAGCTGTATGGAACATCA 726

```
QY 1202 GGAAGAAAGCACAGGGGCC 1220
Db 727 GCACGAAACACAGGGGGC 745

RESULT 10
BG758425 843 bp mRNA linear EST 15-MAY-2001
602712623F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853164 5',
mRNA sequence.
ACCESSION BG758425
VERSION BG758425.1 GI:14069078
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 843)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1698 row: m column: 05
High quality sequence stop: 835.
Location/Qualifiers
1. 843
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:4853164"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 48"
/notes="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 31.5%; Score 600.6; DB 4; Length 843;
Best Local Similarity 96.2%; Pred. No. 2.4e-151;
Matches 659; Conservative 0; Mismatches 19; Indels 7; Gaps 4;

QY 544 AGCACTTTTATGAAGACTGCTTTTGTGATGATGAAGAAGTCCAGTATGCTTCCT 603
Db 142 AGTACTTTTATGAAGACTGCTTTTGTGATGATGAAGAAGTCCAGTATGCTTCCT 201

QY 604 ACCATGGCAGCAGGTCTGAATCCATCTTTTGGGATTAACATCGACAAACAGGATTTG 663
Db 202 ACCATGGCAGCAGGTCCGAATCCATCTTTTGGGATTAACATCGACAAACAGGATTTG 261

QY 664 AACGGGCAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTTTAAAGGAGTCAACCGAAGC 723
Db 262 AACGGGCAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTTTAAAGGAGTCAACCGAAGT 321

QY 724 GTGACCTCTTGTGGAAGGAGTCCACGCAAGGAGTGAAGCAGCTGTTTCAGGGAGATCACA 783
Db 322 GTGA---CCTTGTCTGAAGGAGTCCACGCAAGGAGTGAAGCAGCTGTTTCAGGGAGATCACA 378

QY 784 GCCTTCCTCTGCGCTCTCATCTCATCAAACTGAACAGGAGACCGA-CCCTTCCTGTC 842
```

```
|||||
379 GCCTTCCTCTGCCATCTCCATCTCATCAAACTGAACAGGAGACCGACCCCTTGCCCGTC 438
QY 843 GTGTCCAGGAATGTCAGTGTGATGCCAAATGCAAAAGGAGCGGAAGAGAAAGAAA 902
Db 439 GTGTCCAGGAATGTCAGTGTGATGCCAAATGCAAAAGGAGCGGAAGAGAAAGAAA 498
QY 903 GTGACCAACAATACTCTATTGTGATGAGGAAGATGAGCAGAACTCTCGGCGAGCTGTTT 962
Db 499 GTGACCAACAATACTCTATTGTGATGAGGAAGATGAGCAGAACTCTCGGCGAGCTGTTT 558
QY 963 AAAAAGACACCTGGGGCAGGGAGAGTCTCAGAGACAACTCCCGACCGCTCTCTGTCAAT 1022
Db 559 AAAAAGACACCTGGGGCAGGGAGAGTCTCAGAGACAACTCCCGACCGCTCTCTGTCAAT 618
QY 1023 ATCATGTCCGCTTTGAAAGCCCTTCGGGCGCTTAATCCCAATGGA-TCAGAGCAGCAAC 1081
Db 619 ATCATGTCCGCTTTGAAAGCCCTTCGGGCGCTTAATCCCAATGGAAGTTCAGAGCAGCAAC 678
QY 1082 TCATGG--AAAATTGATTCCTGTCTTTGAAACGGGGAGTTTGGGTACGAAAGCTTGATG 1139
Db 679 TCGTGGGAAAATTGGATTCCTGTCTTTGAAACAGGGAGTTTGGGTACGAAAGCTTGATG 738
QY 1140 TGAAGAAGCATCGATGATGAAGATGTGATGAAACGAAGATGACGTGTATGGAACATCAT 1199
Db 739 TGAAGAAGCATCGATGATGAAGATGTGATGAAACGAAGATGACGTGTATGGAACATCAT 798
QY 1200 CAGGAAGGAAAGCAGAGGGGCCATC 1224
Db 799 CAGGACGGAAGACAGAGGGGCCATC 823

RESULT 11
BQ057008 892 bp mRNA linear EST 29-MAR-2002
LOCUS AGENCOURT_6769190 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5812156
DEFINITION 5', mRNA sequence.
ACCESSION BQ057008
VERSION BQ057008.1 GI:19816348
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 892)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2062 row: c column: 05
High quality sequence stop: 608.
Location/Qualifiers
1. 892
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5812156"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
```

of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

ORIGIN		Query Match		31.5%; Score 600.4; DB 5; Length 892;		Best Local Similarity		92.1%; Pred. No. 2.8e-151;		Matches 678; Conservative 0; Mismatches 51; Indels 7; Gaps 4;	
QY	544	AGCACATTTTATGAAGACTGGTCTTTGTGATGGATGAAGAGTCCAGTATGCTTCT	603								
Db	98	AGTACTTTTATGAAGACTGGTCTTTGTGATGGATGAAGAGTCCAGTATGCTTCT	157								
QY	604	ACCATGGCAGCAGTCTGAACCTCATACTCTTTGCGATTAAACATCGACAAACAGGATTG	663								
Db	158	ACCATGGCAGCAGTCCGAACCTCATACTCTTTGCGATTAAACATCGACAAACAGGATTG	217								
QY	664	AACGGCAGAGTAAGTTTGTCTCCACGTTTCAGACTCTTAAAGAGTCAACGAGAAC	723								
Db	218	AACGGCAGAGTAAGTTTGTCTCCACGTTTCAGACTCTTAAAGAGTCAACGAGAAC	277								
QY	724	GTGACCTCTTGTGTAAGGAGTCCACGCAAGGAGTGCAGCAGCTGTTTCAGGGAGATCA	783								
Db	278	GTGAC---CTTGCTGAAGGAGTCCACGCAAGGAGTGCAGCAGCTGTTTCAGGGAGATCA	334								
QY	784	GCCTCTCTGCCCTCTCCATCTCATCAACCTGAACAGGAGACCGA-CCCTTGCTCTGTC	842								
Db	335	GCCTCTCTGCCATCTCCATCTCATCAACCTGAACAGGAGACCGACCCCTTGCCGTC	394								
QY	843	GTGTCAGGAATGTGAGTGTGATGTCATCAATGCAAAAGAGCGGAAGAAAGAA	902								
Db	395	GTGTCAGGAATGTGAGTGTGATGTCATCAATGCAAAAGAGCGGAAGAAAGAA	454								
QY	903	GTGACCAACATATCTCATTTGATGATGAGGAGTGCAGCAACTCTGGGACGTTT	962								
Db	455	GTGACCAACATATCTCATTTGATGATGAGGAGTGCAGCAACTCTGGGACGATGTT	514								
QY	963	AAAAAGACACCTGGGCGAGGAGAGCTCAGAGCAACTCCGACCGCTCTCTGTCAAT	1022								
Db	515	AAAAAGACACCTGGGCGAGGAGAGCTCAGAGCAACTCCGACCACTCTCTGTCAAT	574								
QY	1023	ATCATGTCGCTTTGAAAGCCCTTTCGGGCTTAATCCAAATGGAA-TCAGAGCAGCAAC	1081								
Db	575	ATCATGTCGCTTTGAAAGCCCTTTCGGGCTTAATCCAAATGGAAAGTCAGAGCAGCAAC	634								
QY	1082	TCATGAAATATGATTCCTCTGTTTGAACGGGAGTTTGGGTACCAAGCTTGATGTG	1141								
Db	635	TCGTGAAATATGATTCCTCTGTTTGAACAGGGAGTTTGGGTACCAAGCTTGATGTG	694								
QY	1142	AAAAAGCATCGATGATGATGGAATGAAACGAAGATGACGTGATGGAACCTCATCA	1201								
Db	695	AAAAAGCATCGATGATGGAATGGAATGAAACGAAGATGACGTGATGGAACCTCATCA	754								
QY	1202	GGAA--GGAAGCACAGGGGCCACTCGGAGTCGCGGAGAGCCACTTGGGAAGGGAA	1259								
Db	755	GGAGCGAAGGCCAGGGGCCACTCGGAGATCCCGCGAGAAAACGGAACACACTTCT	814								
QY	1260	GCCTCTCCAGATGCA	1275								
Db	815	GTACACCCAGGCTGGA	830								

RESULT 12
BM692924
LOCUS
DEFINITION
UI-E-CK1-afk-h-18-0-UI.r1 UI-E-CK1 Homo sapiens cDNA clone
UI-E-CK1-afk-h-18-0-UI 5', mRNA sequence.
ACCESSION
BM692924
VERSION
BM692924.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 609)
Bonaldo, M.P., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse

FEATURES
source

Location/Qualifiers
1..609
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CK1-afk-h-18-0-UI"
/tissue_type="Retina Foveal and Macular"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CK1"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-CK1 is a normalized cDNA library containing the
following tissue(s): Retina Foveal and Macular. The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is GTCC. This library was created for the program,
Gene Discovery in the Visual System, supported by National
Eye Institute (NEI)."

ORIGIN

Query Match		30.7%; Score 586; DB 4; Length 609;		Best Local Similarity		99.7%; Pred. No. 2e-147;		Matches 608; Conservative 0; Mismatches 0; Indels 2; Gaps 2;	
QY	1215	GGGGCCACTCGGAGTCGCCGAGAGCCACTGGAGGGAAACACCTGCCCTCTCCAGATGC	1274						
Db	1	GGGGCCACTCGGAGTCGCCGAGAGCCACTGGAGGGAAACACCTGCCCTCTCCAGATGC	60						
QY	1275	ACAGCTGGGCTCCGCTGAAGGTGCTGCAATGATCTCGACATCTCTTCCCTGTCACTG	1334						
Db	61	ACAGCTGGGCTCCGCTGAAGGTGCTGCAATGATCTCGACATCTCTTCCCTGTCACTG	120						
QY	1335	GGTGGGCTCTACAGCCGACAGATGCCCTCGGAGCCTGGAGACGGACAGGAC	1394						
Db	121	GGTGGGCTCTACAGCCGACAGATGCCCTCGGAGCCTGGAGACGGACAGGAC	180						
QY	1395	CAGAGACACCGTTCTCCCGACTCTCGGTACAGTCTGGAAGCCAGCTCTCCAG	1454						
Db	181	CAGAGACACCGTTCTCCCGACTCTCGGTACAGTCTGGAAGCCAGCTCTCCAG	240						
QY	1455	GCCACGGAAGTCTCTGAGCAGCCTGTGA-CTTCTGCTCTAGTGCAGAGTCCATGACAA	1513						

```
Db 241 GCCAGGAAGTCTCTGAGCAGCCTGTTACCTTCTGCTCAGTCCAGATCCATGACAA 300
Qy 1514 TTAGTCACTGCGCAGGCCACTGTGGCCATCATGAACAGAGAGATAGCTGGAGGAGG 1573
Db 301 TTAGTGAACCTGCGCAGGCCACTGTGGCCATCATGAACAGAGAGATAGCTGGAGGAGG 360
Qy 1574 AGAAGAGATCACTGCGAAACCTGCTCGACGGTGAGATGAGCACTCAGCCGGCTCCGGC 1633
Db 361 AGAAGAGATCACTGCGAAACCTGCTCGACGGTGAGATGAGCACTCAGCCGGCTCCGGC 420
Qy 1634 AAGAGGTGACACCTTTGAAAAGAGGTGGCTGAAACAGAGAGAGCGGAGGCGCATGAAGG 1693
Db 421 AAGAGGTGACACCTTTGAAAAGAGGTGGCTGAAACAGAGAGAGCGGAGGCGCATGAAGG 480
Qy 1694 TCCAGCGCTGCGCAGCTATCTTTGCTATTTTGTGAGGAGATCTTAAACCCACAGTGAGAA 1753
Db 481 TCCAGCGCTGCGCAGCTATCTTTGCTATTTTGTGAGGAGATCTTAAACCCACAGTGAGAA 540
Qy 1754 CCATGTGTGGAGAAATGGAGGAGAGAGAAATCCAAACAGTTCTCTGATAGTCTCATTTGA 1813
Db 541 CCATGTGTGGAGAAATGGAGGAGAGAGAAAT-CAACAGTTCTCTGATAGTCTCATTTGA 599
Qy 1814 GCTCTCTGAT 1823
Db 600 GCTCTCTGAT 609
```

RESULT 13

```
LOCUS BG684262 813 bp mRNA linear EST 01-MAY-2001
DEFINITION 602635896F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4763783 5',
mRNA sequence.
ACCESSION BG684262
VERSION BG684262.1 GI:13915659
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 813)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM1619 row: p column: 24
High quality sequence stop: 795.
Location/Qualifiers
1. .813
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4763783"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 48"
/notes="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >50bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
```

FEATURES

```
source
1. .813
BQ184171/c 626 bp mRNA linear EST 30-APR-2002
LOCUS UI-E-EJ1-ajs-e-23-0-UI.s1 UI-E-EJ1 Homo sapiens cDNA clone
DEFINITION UI-E-EJ1-ajs-e-23-0-UI 3', mRNA sequence.
ACCESSION BQ184171
VERSION BQ184171.1 GI:20359722
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 626)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovey
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
```

ORIGIN

Db	177	AGTACTTTTATGAAGACTCGTCTTTTGTGTGATGGATGAAGAGAGGTCCTAGTAGCTTCTCT	236
Qy	604	ACCATGGCAGCAGGTCTGAACTCCATACTCTTTGGCATTAACATCGACAAACGAAGATTG	663
Db	237	ACCATGGCAGCAGGTCCGAATCCATACCTTTGGCATTAACATTCGACAAACGAAGATTG	296
Qy	664	AACGGGCGAGAGTAAGTTTGTCTCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCCAGAAC	723
Db	297	AACGGGCGAGAGTAAGTTTGTCTCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCCAGAAT	356
Qy	724	GTGACCTCTCTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTTCAGGGAGATCACA	783
Db	357	GTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTTCAGGGAGATCACA	413
Qy	784	GCCTCTCTGCGGTCTCATCTCTCATCAAACTTGAACAGGAGACCGA-CCCTTGCCTGTCTC	842
Db	414	GCCTCTCTGCGCATCTCATCTCTCATCAAACTTGAACAGGAGACCGACCCCTTGCCTGTCTC	473
Qy	843	GTGTCCAGGAATGTCAAGTGTGATGTCMAATGCMAAAGAGGAGCGGAAGAAAGAAAAAGAA	902
Db	474	GTGTCCAGGAATGTCAAGTGTGATGTCMAATGCMAAAGAGGAGCGGAAGAAAGAAAAAGCAA	533
Qy	903	GTGACCAACATAATCTCATTTTGTATGATGAGGAAGATGAGCAGAACTCTGGGACGTGTTT	962
Db	534	GTGACCAACATAATCTCATTTTGTATGATGAGGAAGATGAGCAGAACTCTGGGACATGTTT	593
Qy	963	-AAAAAGACACTGGGGCAGGGGAGAGCTCAGAGGAACAACTCCGACCGCTCTCTGTCAA	1021
Db	594	AAAAAAGAACCTGGGGCAGGGGAGAGCTCAGAGGAACAACTCCGACCACTCTCTGTCAA	653
Qy	1022	TATCATGTCCGCTTTGAAAGCCCTTCGGGCCTAACTCCAATGGATCAGAGCAGCAAC	1081
Db	654	TATCATGTCCGCTTTGAAAGCCCTTCGGGCCTAACTCCAATGGAAAGTCAGAGCAGCAAC	713
Qy	1082	TC-ATGGAAAAATTGATTCCTGTCTC--TTTGAAACGGGGAGTTTGGGTACCGAAGCTTGAT	1138
Db	714	TCGTTGGAAAAATTGATTCCTGTCTCCTTTGAACAGGGGAGTCTGGGTACCGAAGCTTGAT	773
Qy	1139	GTGAAAAAGCATCGATGATGAAGATGT--GGATGAAAAACGAAGTGA	1182
Db	774	GTGAAAAAGCATCGATGATGAAGATGTGGGATGAAAAACGAAGATTA	819

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	133	4	7.0	190	4	US-09-513-999C-16002
2	83	6	4.4	100	4	US-09-513-999C-16212
3	66	8	3.5	811	4	US-09-270-767-12402
C 4	55	6	2.9	7218	1	US-08-232-463-14
	54	8	2.9	36731	4	US-09-949-016-13770
C 5	54	8	2.9	601	4	US-09-949-016-69138
	54	4	2.9	54	4	US-09-573-080A-118
C 6	53	2	2.8	476	4	US-09-976-594-221
C 7	52	2	2.7	961	4	US-09-949-016-12759
	50	2.6	9811	4	US-09-949-016-15004	
9	46	8	2.5	60276	4	US-09-949-016-15694
C 10	46	8	2.5	60338	4	US-09-949-016-11901
C 11	46	8	2.5	94133	4	US-09-949-016-12713
	46	8	2.5	94135	4	US-09-949-016-15934
C 12	46	8	2.5	94135	4	US-09-949-016-15936
C 13	46	8	2.5	94135	4	US-09-165-827C-1
C 14	46	8	2.5	94135	4	US-09-165-827C-13
C 15	46	8	2.5	94135	4	US-09-621-976-2453
C 16	46	8	2.5	94135	4	US-09-621-976-2453
C 17	46	8	2.5	94135	4	US-09-506-066E-9
C 18	46	6	2.4	1182	3	US-09-621-976-18033
19	46	6	2.4	1185	3	US-09-498-557-20
20	45	6	2.4	551	4	US-09-949-016-12107
21	45	6	2.4	571	4	US-09-949-016-15878
22	45	4	2.4	1307	4	US-09-949-016-12147
23	44	8	2.4	474	4	US-09-949-016-12147
24	44	8	2.4	2056	4	US-09-949-016-12107
C 25	44	6	2.3	23802	4	US-09-949-016-12107
C 26	44	6	2.3	23803	4	US-09-949-016-12147
C 27	43	8	2.3	767577	4	US-09-949-016-12147

[illegible]

Qy	1060	CAATGGAAATCAGACGCAACTCATGGAATAATTGATTCCTCGTCTTTTGAAACGGGAGTT	1119
Db	1256	RR	1197
Qy	1120	TGGGTACAGAAAGCTTGATGTGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGA	1179
Db	1196	RR	1137
Qy	1180	TGACGTGTATGAAACTCATCAGGAAGAACACACAGGGGCCACTCGAGTCGCCCGAGAA	1239
Db	1136	RR	1077
Qy	1240	GCCACTGGAAGGGAACACTGCCTCTCCAGATGCACAGCTGGG	1283
Db	1076	RRRRRRRRRRATCGCAAGCTCCCTGACCTGCAGCCAAAGCTCGG	1033

RESULT 5
 US-09-949-016-13770/c
 ; Sequence 13770, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13770
 ; LENGTH: 36731
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-13770

Query Match	2.9%	Score 54.8;	DB 4;	Length 36731;
Best Local Similarity	63.8%;	Pred. No. 0.00062;		
Matches 83;	Conservative 0;	Mismatches 47;	Indels 0;	Gaps 0;

Qy	1768	AATGAGGGAGAGAGAAATCAACAGTTCCTGTGATAGTCTCATTTGAGCTCTCGATCCAG	1827
Db	6869	AAAAGTTGAAACCAAGACGGAGAAAGATTCTGTGATGACCTATTAGATAACCTGGATCCAG	6810
Qy	1828	TCTTTCCTGAAGCTGTGTTTCTCTGGACTTTTCATGTATGTGACCAATAAATTCCTTT	1887
Db	6809	CCTTGCCUGAATGACGATATACCTCTGGATTCTTAGTCATGTGAGCCAAAGTATGATCCGCTG	6750
Qy	1888	CATTCCTTGA	1897
Db	6749	TTTTGCTTAA	6740

RESULT 6
 US-09-949-016-69138/c
 ; Sequence 69138, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768

QY 1890 TTCCTTGA 1897
|||
Db 433 TTGCTTAA 440
|||
RESULT 8
US-09-976-594-221
; Sequence 221, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 221
; LENGTH: 961
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 330927.1
US-09-976-594-221
Query Match 2.7%; Score 52.2; DB 4; Length 961;
Best Local Similarity 71.1%; Pred. No. 0.00052;
Matches 69; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 1793 GTTCTGATGTCATTTAGAGTCCTGGATCCAGTCCTTCTCTGAAGTGTGTTCTCT 1852
|||
Db 861 GTCTTGTGTCATCTTTTGAGCCACTAGACCAAGCTTTACCTGAAGCAGAGCTACCTCA 920
|||
QY 1853 GGACTTTTCATGTATGTGAGCCCAATAATTGCTTTCA 1889
|||
Db 921 GAATTTTCAGCTATGTGAGCCCAATAACATCTGTCA 957
|||
RESULT 9
US-09-949-016-12759
; Sequence 12759, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12759
; LENGTH: 9811
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12759
Query Match 2.6%; Score 50; DB 4; Length 9811;
Best Local Similarity 61.5%; Pred. No. 0.0067;
Matches 80; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 1768 AATGAGGGAGAGAGAAATCCACAGTTCCTGTAGTCTCATTGAGCTCCTGGATCCAG 1827

Db 3752 AAGAAAGAGAGTGAGTGAGTGAGTCCATTTCTTTATTCATCTCTTTGAGGCCCTACACCTAA 3811
|||
QY 1828 TCCTTCTGAGAGCTGTGTTCCCTCTGAGCTTTTCATGATGTAGGCAATAAATGCTTT 1887
|||
Db 3812 TCATTGATGAAGCTATATACCACTCCAGCCTTTTCAATTACTTGAGCCAATCAATTCCTT 3871
|||
QY 1888 CATTCTTTGA 1897
|||
Db 3872 TAATAGTTAA 3881
|||
RESULT 10
US-09-949-016-15004/c
; Sequence 15004, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15004
; LENGTH: 60276
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(60276)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15004
Query Match 2.5%; Score 46.8; DB 4; Length 60276;
Best Local Similarity 73.2%; Pred. No. 0.13;
Matches 60; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 1806 TCATTGAGCTCCTGGATCCAGTCTTTCTCTGAAGCTGTGTTCTCTGAGCTTTTCATGT 1865
|||
Db 10119 TCTTAGAAGCACCTGGATCCAGCCCTGCTGAGCCATCTTACCCCTGAGACTTCTCAGTT 10060
|||
QY 1866 ATGTGAGCCCAATAAATGCTTT 1887
|||
Db 10059 GTGTGAGCCCAAGAATTCCTTT 10038
|||
RESULT 11
US-09-949-016-15694/c
; Sequence 15694, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 15694
; LENGTH: 60338
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(60338)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15694

Query Match
Best Local Similarity 2.5%; Score 46.8; DB 4; Length 60338;
Matches 60; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1806 TCATTGAGCTCTGATCAGTCCTTCTGAGCTGTTCTCTGAGCTTTTCATCT 1865
Db 10181 TCTTAGAGCACCCTGGATCAGCCCTGCTGAGCCATCTACCCCTGGACTTCTCAGTT 10122

QY 1866 ATGTGAGCCCAATAAATGCTTT 1887
Db 10121 GTGTGAGCAAGAATGTCTTT 10100

RESULT 12
US-09-949-016-11901/c
; Sequence 11901, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11901
; LENGTH: 94133
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11901

Query Match
Best Local Similarity 2.5%; Score 46.8; DB 4; Length 94133;
Matches 66; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1792 AGTTCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCTGAGCTGTGTTCCCTC 1851
Db 40227 AGAGCCTGAGACACAGTTTAAGCATTTGGAGTCAGCTATGCTGAGCCATGCTCCCTC 40168

QY 1852 TGGACTTTTCATGATGTGAGCCCAATAAATGCTTTCA 1889
Db 40167 TAGACTTTTCAGTTACATGAGCCCAATAAATATTTTA 40130

RESULT 13
US-09-949-016-12713/c
; Sequence 12713, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12713
; LENGTH: 94133
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12713

Query Match
Best Local Similarity 2.5%; Score 46.8; DB 4; Length 94133;
Matches 66; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1792 AGTTCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCTGAGCTGTGTTCCCTC 1851
Db 40227 AGAGCCTGAGACACAGTTTAAGCATTTGGAGTCAGCTATGCTGAGCCATGCTCCCTC 40168

QY 1852 TGGACTTTTCATGATGTGAGCCCAATAAATGCTTTCA 1889
Db 40167 TAGACTTTTCAGTTACATGAGCCCAATAAATATTTTA 40130

RESULT 14
US-09-949-016-15934/c
; Sequence 15934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15934
; LENGTH: 94135
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15934

Query Match
Best Local Similarity 2.5%; Score 46.8; DB 4; Length 94135;
Matches 66; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1792 AGTTCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCTGAGCTGTGTTCCCTC 1851
Db 40227 AGAGCCTGAGACACAGTTTAAGCATTTGGAGTCAGCTATGCTGAGCCATGCTCCCTC 40168

QY 1852 TGGACTTTTCATGATGTGAGCCCAATAAATGCTTTCA 1889
Db 40167 TAGACTTTTCAGTTACATGAGCCCAATAAATATTTTA 40130

RESULT 15
US-09-949-016-15935/c
; Sequence 15935, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

Mon Jun 13 13:27:09 2005

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15935
; LENGTH: 94135
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15935

Query Match      2.5%; Score 46.8; DB 4; Length 94135;
Best Local Similarity 67.3%; Pred. No. 0.16; 32; Indels 0; Gaps 0;
Matches 66; Conservative 0; Mismatches 0;

QY 1792 AGTTCCTGATAGTCTCAATTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTCCCTC 1851
Db 40227 AGAGCCTGAGAACACACAGTTTAAGCATTTGGAGTCAGCTATGCCTGAAGCCATGCTCCCTC 40168
QY 1852 TGGACTTTTCATGTATGTGAGCCATAAATTGCTTCA 1889
Db 40167 TAGACTTTTCAGTTACATGAGCCATAAAATATTTT 40130

Search completed: June 10, 2005, 20:46:32
Job time : 213.376 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 06:13:19 ; Search time 4767.2 Seconds
(without alignments)
16321.674 Million cell updates/sec

Title: US-09-155-676B-1

Perfect score: 1906

Sequence: 1 cattgggtcacgcgggtggcg.....tccattccttgaaaaaaa 1906

Scoring table: IDENTITY NUC

Gapop 10'0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCTUS1 COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCTUS2 COMB.seq:*
3: /cgn2_6/ptodata/1/pna/PCTUS3 COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US06 COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US07 COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US081 COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US082 COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US083 COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US084 COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US085 COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US086 COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US087 COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US088 COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US089 COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US090 COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US091 COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US092A COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US092B COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US093A COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US093B COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US094 COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US095A COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US095B COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095C COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US095D COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096A COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096B COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096C COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096D COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US096E COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097A COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US097B COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US097C COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US098A COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US098B COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US098C COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US098D COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099A COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US099B COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US099C COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US099D COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US099E COMB.seq:*

/cgn2_6/ptodata/1/pna/US099F COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US099G COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US100A COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US100B COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US100C COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US101A COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US101B COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US102A COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US102B COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US103A COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US103B COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US104A COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US104B COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US105A COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US105B COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US106A COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107A COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107B COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107C COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US107D COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US108A COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US108B COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US109A COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US109B COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US109C COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US110 COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6000 COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6001 COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6002 COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6003 COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6004 COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6005 COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6006 COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6007 COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6008 COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6009 COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6010 COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6011 COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6012 COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6013 COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6014 COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6015 COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6016 COMB.seq:*
86: /cgn2_6/ptodata/1/pna/US6017 COMB.seq:*
87: /cgn2_6/ptodata/1/pna/US6018 COMB.seq:*
88: /cgn2_6/ptodata/1/pna/US6019 COMB.seq:*
89: /cgn2_6/ptodata/1/pna/US6020 COMB.seq:*
90: /cgn2_6/ptodata/1/pna/US6021 COMB.seq:*
91: /cgn2_6/ptodata/1/pna/US6022 COMB.seq:*
92: /cgn2_6/ptodata/1/pna/US6023A COMB.seq:*
93: /cgn2_6/ptodata/1/pna/US6023B COMB.seq:*
94: /cgn2_6/ptodata/1/pna/US6024 COMB.seq:*
95: /cgn2_6/ptodata/1/pna/US6025 COMB.seq:*
96: /cgn2_6/ptodata/1/pna/US6026 COMB.seq:*
97: /cgn2_6/ptodata/1/pna/US6027 COMB.seq:*
98: /cgn2_6/ptodata/1/pna/US6028 COMB.seq:*
99: /cgn2_6/ptodata/1/pna/US6029 COMB.seq:*
100: /cgn2_6/ptodata/1/pna/US6030 COMB.seq:*
101: /cgn2_6/ptodata/1/pna/US6031 COMB.seq:*
102: /cgn2_6/ptodata/1/pna/US6032 COMB.seq:*
103: /cgn2_6/ptodata/1/pna/US6033 COMB.seq:*
104: /cgn2_6/ptodata/1/pna/US6034 COMB.seq:*
105: /cgn2_6/ptodata/1/pna/US6035 COMB.seq:*
106: /cgn2_6/ptodata/1/pna/US6036 COMB.seq:*
107: /cgn2_6/ptodata/1/pna/US6037 COMB.seq:*
108: /cgn2_6/ptodata/1/pna/US6038 COMB.seq:*
109: /cgn2_6/ptodata/1/pna/US6039 COMB.seq:*
110: /cgn2_6/ptodata/1/pna/US6040 COMB.seq:*
111: /cgn2_6/ptodata/1/pna/US6041 COMB.seq:*
112: /cgn2_6/ptodata/1/pna/US6042 COMB.seq:*
113: /cgn2_6/ptodata/1/pna/US6043 COMB.seq:*
114: /cgn2_6/ptodata/1/pna/US6044 COMB.seq:*
115: /cgn2_6/ptodata/1/pna/US6045 COMB.seq:*
116: /cgn2_6/ptodata/1/pna/US6046 COMB.seq:*

```

RESULT 1
US-09-155-676-1
; Sequence 1, Application US/09155676
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSER: BROWDY AND NEIMAR, P.L.L.C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/155,676
; FILING DATE: 04-JAN-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IL97/00117
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117800
; FILING DATE: 02-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 119133
; FILING DATE: 26-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: WALLACH=21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1906 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-155-676-1

Query Match 99.6%; Score 1899; DB 17; Length 1906;
Best Local Similarity 100.0%; Pred. No. 0; Indels 0; Gaps 0;
Matches 1906; Conservative 0; Mismatches 0;

Qy 1 CATTGGGTACGCGGTGGCGCGCTCTAGAAATAGTGGATCCCCGGGCTGCAGGAATTGC 60
Db 1 CATTGGGTACGCGGTGGCGCGCTCTAGAAATAGTGGATCCCCGGGCTGCAGGAATTGC 60
Qy 61 ATTTCGAGGCCACGAAGCCGCGCGCGCCANGCACCGCCCGGGANAGGNCCCATG 120
Db 61 ATTTCGAGGCCACGAAGCCGCGCGCGCCANGCACCGCCCGGGANAGGNCCCATG 120
Qy 121 AGCGGATCNCNGAACNATGACAAAGACAAATTTCTGCTGGAGCGACTGCTGGATGCATG 180
Db 121 AGCGGATCNCNGAACNATGACAAAGACAAATTTCTGCTGGAGCGACTGCTGGATGCATG 180
Qy 181 AAAACAGTGCAGATCCGCTTTTNGAGGGAGAAAGAGATTGCTCGGATTCGACAGCAGG 240
Db 181 AAAACAGTGCAGATCCGCTTTTNGAGGGAGAAAGAGATTGCTCGGATTCGACAGCAGG 240

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	1899	99.6	1906	17	US-09-155-676-1	Sequence 1, Appli	
2	1899	99.6	1906	17	US-09-155-676A-1	Sequence 1, Appli	
3	1742	91.4	1782	47	US-10-070-255-4	Sequence 4, Appli	
4	1555	81.6	2873	47	US-10-070-255-6	Sequence 5, Appli	
5	1555	81.6	3139	47	US-10-070-255-5	Sequence 6, Appli	
6	1138.2	59.7	2248	28	US-09-629-465A-14016	Sequence 14016, A	
7	1138.2	59.7	2248	65	US-10-917-503-14016	Sequence 14016, A	
8	1032.4	57.3	1516	22	US-09-471-275-4362	Sequence 4362, Ap	
9	1092.4	57.3	1516	26	US-09-488-725B-9141	Sequence 9141, Ap	
10	1092.4	57.3	1516	26	US-09-552-317-9141	Sequence 9141, Ap	
11	883.6	46.4	1137	33	US-09-758-472-1043	Sequence 1043, Ap	
12	883.6	46.4	1137	50	US-10-235-926-1043	Sequence 1043, Ap	
13	792.8	41.6	6045	28	US-09-629-469A-18174	Sequence 18174, A	
14	792.8	41.6	6045	65	US-10-917-503-18174	Sequence 18174, A	
15	777.4	40.8	1630	90	US-60-212-356-340	Sequence 340, App	
16	774.2	40.6	1642	90	US-60-212-356-343	Sequence 373, App	
17	765.6	40.2	814	49	US-10-170-235-20448	Sequence 20448, A	
18	738.8	38.8	1147	48	US-10-144-771-19120	Sequence 19120, A	
19	738.8	38.8	1147	106	US-60-360-207-19120	Sequence 19120, A	
20	719.8	37.8	834	21	US-09-396-087-4235	Sequence 4235, Ap	
21	672.2	35.3	730	28	US-09-629-469A-1808	Sequence 1808, Ap	
22	672.2	35.3	730	65	US-10-917-503-1808	Sequence 1808, Ap	
23	565.4	29.7	575	25	US-09-540-212A-52972	Sequence 52972, A	
24	554.4	29.1	576	97	US-60-278-561-628	Sequence 628, App	
25	535	28.1	542	25	US-09-540-229-98596	Sequence 98596, A	
26	456.4	23.9	512	21	US-09-396-087-160	Sequence 160, App	
27	446.2	23.4	554	43	US-10-170-235-33933	Sequence 33933, A	
28	440.4	23.1	1667	87	US-60-185-213-1385	Sequence 1385, Ap	
29	433.6	22.7	506	25	US-09-540-212A-44800	Sequence 44800, A	
30	421.8	22.1	437	21	US-09-362-510-18241	Sequence 18241, A	
31	421.8	22.1	437	21	US-09-363-510A-18241	Sequence 18241, A	
32	421.8	22.1	437	39	US-09-004-013-18241	Sequence 18241, A	
33	417.4	21.9	2075	91	US-60-229-525-567	Sequence 567, App	
34	412.4	21.6	422	23	US-09-528-403-108313	Sequence 108313, A	
35	412.4	21.6	422	40	US-09-933-524-108313	Sequence 108313, A	
36	412.4	21.6	422	40	US-09-933-524A-108313	Sequence 108313, A	
37	404.6	21.2	455	23	US-09-528-409-108118	Sequence 108118, A	
38	404.6	21.2	455	40	US-09-933-524-108118	Sequence 108118, A	
39	404.6	21.2	455	49	US-09-933-524A-108118	Sequence 108118, A	
40	393	20.6	1547	49	US-10-170-235-6644	Sequence 6644, Ap	
41	388.2	20.4	406	19	US-09-287-618-7448	Sequence 7448, Ap	
42	382	20.0	422	23	US-09-528-409-78923	Sequence 78923, A	
43	382	20.0	422	40	US-09-933-524-78923	Sequence 78923, A	
44	382	20.0	422	40	US-09-933-524A-78923	Sequence 78923, A	
45	374.4	19.6	391	97	US-60-378-561-827	Sequence 827, App	

ALIGNMENTS

241 GTACCTGTCTGTGCCCCAGTTTGAAGCCGTCTCTGAGCATGGCTTGAAGAGGAGTCGA 300
Db GTACCTGTCTGTGCCCCAGTTTGAAGCCGTCTCTGAGCATGGCTTGAAGAGGAGTCGA 300
QY GGATTGGCACTCAGCGGCGAGCGATCAAGCAGCGAGCGGGCTTTCGCCAGCAAAACCGAA 360
Db GGATTGGCACTCAGCGGCGAGCGATCAAGCAGCGAGCGGGCTTTCGCCAGCAAAACCGAA 360
QY ACAGAGCCCGTGTCTTGGTACTAGTGAAGAGGTCTCAACAGCAGCAGCTGCGAGCGC 420
Db ACAGAGCCCGTGTCTTGGTACTAGTGAAGAGGTCTCTCAACAGCAGCAGCTGCGAGCGC 420
QY TTCTACTCCCTGCGCCACATCGCCTCAGAGTGGCGGGGTGCGCCCTGGCTGGCTGT 480
Db TTCTACTCCCTGCGCCACATCGCCTCAGAGTGGCGGGGTGCGCCCTGGCTGGCTGT 480
QY GCCTCAACGAACACTCCCTGGAGCGCTACCTGCATGCTCTGCGCCGACCGCTCAGG 540
Db GCCTCAACGAACACTCCCTGGAGCGCTACCTGCATGCTCTGCGCCGACCGCTCAGG 540
QY CTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGTCCAGTATGCTT 600
Db CTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGTCCAGTATGCTT 600
QY CTTACCATGGCAGCGTCTGAACCTCATCTCTTTGGCATTAACATCGACAAACAGGAT 660
Db CTTACCATGGCAGCGTCTGAACCTCATCTCTTTGGCATTAACATCGACAAACAGGAT 660
QY TTGAACGGGCGAGTAGTAACTTGTCTCCACCGCTTTCAGAGCTCTTAAAGGAGTCAACGCG 720
Db TTGAACGGGCGAGTAGTAACTTGTCTCCACCGCTTTCAGAGCTCTTAAAGGAGTCAACGCG 720
QY AACGTGACCTCTTGTGAGGAGTCCAGCAGGAGTGAAGAGCTGTTTCAGGAGATC 780
Db AACGTGACCTCTTGTGAGGAGTCCAGCAGGAGTGAAGAGCTGTTTCAGGAGATC 780
QY ACAGCTCTCTGCGCTCCATCTCTCAATCAAACTCATCAAACTCATCAAACTCATCAAACT 840
Db ACAGCTCTCTGCGCTCCATCTCTCAATCAAACTCATCAAACTCATCAAACTCATCAAACT 840
QY TCCTGTCCAGGATGTCAGTCTGATGCCAAATGCAAAAGGAGCGGAAGCAAGAAAGA 900
Db TCCTGTCCAGGATGTCAGTCTGATGCCAAATGCAAAAGGAGCGGAAGCAAGAAAGA 900
QY AAGTGACCAACATAATCTCATTTGATGATGAGGAGATGAGCAGAACTCTGGGAGCTGT 960
Db AAGTGACCAACATAATCTCATTTGATGATGAGGAGATGAGCAGAACTCTGGGAGCTGT 960
QY TTTAAAGACACTGGGGCAGGGAGAGCTCAGAGCAAACTCCGACCGCTCTCTGTCA 1020
Db TTTAAAGACACTGGGGCAGGGAGAGCTCAGAGCAAACTCCGACCGCTCTCTGTCA 1020
QY ATATCATGTCGCGCTTGAAGCCCTCTGGGCTTAACCTCAATGAATCAGAGCAAA 1080
Db ATATCATGTCGCGCTTGAAGCCCTCTGGGCTTAACCTCAATGAATCAGAGCAAA 1080
QY CTATGGAATAATGATTCCCTGTCTTTGAACGGGAGTTGGGTACAGAGCTTCTATGT 1140
Db CTATGGAATAATGATTCCCTGTCTTTGAACGGGAGTTGGGTACAGAGCTTCTATGT 1140
QY GAAAAGATCGATGATGAAGATGTGATGAAAACGAAGATGAGCTGTATGGAACCTCATC 1200
Db GAAAAGATCGATGATGATGATGAAAACGAAGATGAGCTGTATGGAACCTCATC 1200
QY AGAAGGAAGCACAGGGGCACTCGAGTCCCGGAGAGCCACTGGAAGGAGAACCTTG 1260
Db AGAAGGAAGCACAGGGGCACTCGAGTCCCGGAGAGCCACTGGAAGGAGAACCTTG 1260
QY CCTCTCCAGATCAGAGCTGGGCTCGCTGAAAGTGTCTGCAATGACTCCGACATCCT 1320
Db CCTCTCCAGATCAGAGCTGGGCTCGCTGAAAGTGTCTGCAATGACTCCGACATCCT 1320
QY CTTCCCTGTGAGTGGGCTCTCTACAGCCAGCAGATGCCCTCCCTCGGAAGCCTGGA 1380

1321 CTTCCCTGTGAGTGGGCTCTCTACAGCCAGCAGATGCCCTCGGAAGCCTGGA 1380
QY GAAAGGACAGGACCAAGAGGACCAAGTCTCTCCCGATCTCTGGACTTCGGTACAGTGTGA 1440
Db GAAAGGACAGGACCAAGAGGACCAAGTCTCTCCCGATCTCTGGACTTCGGTACAGTGTGA 1440
QY AGCAGCTCTCAGGCGCAGGAGTCTCTGAGCAGCTGTACTTCTGCTCAGTGCCA 1500
Db AGCAGCTCTCAGGCGCAGGAGTCTCTGAGCAGCTGTACTTCTGCTCAGTGCCA 1500
QY GAGTCATCACAATAGTGAATGCGCCAGGCGCACTGTGGCCATGATGAACAGGAAGGAT 1560
Db GAGTCATCACAATAGTGAATGCGCCAGGCGCACTGTGGCCATGATGAACAGGAAGGAT 1560
QY GAGTCGAGGAGGAGAAACAGATCACTGCGAAACCTCTCGACGGTGAAGGAGCACTCA 1620
Db GAGTCGAGGAGGAGAAACAGATCACTGCGAAACCTCTCGACGGTGAAGGAGCACTCA 1620
QY GCCGCGCTCCGCAAGAGGTGAGACCTTTGAAAAGGAGGTGGCTGAAACAGGAGGCGG 1680
Db GCCGCGCTCCGCAAGAGGTGAGACCTTTGAAAAGGAGGTGGCTGAAACAGGAGGCGG 1680
QY CAGGCGATCAAGGTCCAGGCGCTGGCCAGCTATCTTTGCTATTTCTGAGGAGATCTAA 1740
Db CAGGCGATCAAGGTCCAGGCGCTGGCCAGCTATCTTTGCTATTTCTGAGGAGATCTAA 1740
QY CCCCAGCTCAGAAACCAATGTGGTGGGAAATGAGGAGAGAGAGAAATCCAAACAGTCTCTGA 1800
Db CCCCAGCTCAGAAACCAATGTGGTGGGAAATGAGGAGAGAGAGAAATCCAAACAGTCTCTGA 1800
QY TAGTCTCATTTGAGCTCTGGATCCAGTCTTTCTGAAAGCTGTGTCTCTGAGCTTTT 1860
Db TAGTCTCATTTGAGCTCTGGATCCAGTCTTTCTGAAAGCTGTGTCTCTGAGCTTTT 1860
QY CATGTATGAGCCCAATAAATTGCTTTCACTTCTTGAAAAAATAA 1906
Db CATGTATGAGCCCAATAAATTGCTTTCACTTCTTGAAAAAATAA 1906

RESULT 2

US-09-155-676A-1
; Sequence 1, Application US/09155676A
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/155,676A
; FILING DATE: 04-JAN-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IL97/00117
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117800

FILING DATE: 02-APR-1996
PRIOR APPLICATION DATA: IL 119133
APPLICATION NUMBER: 26-AUG-1996
FILING DATE: 26-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH=21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1906 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-155-676A-1

Query Match 99.6%; Score 1899; DB 17; Length 1906;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1906; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATTGGGTACCGGTGGCGGGCTCTAGAATAGTGGATCCCGGGCTGCAGGAATCG 60
DB 1 CATTGGGTACCGGTGGCGGGCTCTAGAATAGTGGATCCCGGGCTGCAGGAATCG 60

QY 61 ATTCCAGGCCACGAAGCCGGCGCGCGCGCANGCACCGGCCGGGGANAGGCCCATG 120
DB 61 ATTCCAGGCCACGAAGCCGGCGCGCGCGCANGCACCGGCCGGGGANAGGCCCATG 120

QY 121 AGCGGATCNCNGAANCATGACAAAAAGCAATTTCTGTGTGGAGGACTCTCGATGCAATG 180
DB 121 AGCGGATCNCNGAANCATGACAAAAAGCAATTTCTGTGTGGAGGACTCTCGATGCAATG 180

QY 181 AAACAGTGCAGATCCGCTTTNGAGGAGGAAGAGATTGCTCGGATTCCGACAGCAGG 240
DB 181 AAACAGTGCAGATCCGCTTTNGAGGAGGAAGAGAGATTGCTCGGATTCCGACAGCAGG 240

QY 241 GTCACTGTCTGTGTGCCAGTTTGAAGCCGTCTTGCAGCATGGCTTGAAGAGGAGTGA 300
DB 241 GTCACTGTCTGTGTGCCAGTTTGAAGCCGTCTTGCAGCATGGCTTGAAGAGGAGTGA 300

QY 301 GGATTGGCACTACAGCGCGAGGATCAAGCAGGCGGGCTTTGCCAGCAAAACCGAA 360
DB 301 GGATTGGCACTACAGCGCGAGGATCAAGCAGGCGGGCTTTGCCAGCAAAACCGAA 360

QY 361 ACAGAGCCGTGTCTGTGTACTACGTGAAGGAGGTCTTCAACAGCAGAGCTGCAGCG 420
DB 361 ACAGAGCCGTGTCTGTGTACTACGTGAAGGAGGTCTTCAACAGCAGAGCTGCAGCG 420

QY 421 TTCTACTCTCCGCGCACATTCGCTCAGAGTGGGGCGGGGTGCGGCTGGCTGGCTGT 480
DB 421 TTCTACTCTCCGCGCACATTCGCTCAGAGTGGGGCGGGGTGCGGCTGGCTGGCTGT 480

QY 481 GCCTCACAACCACTCCCTGGAGCGCTACCTGCACATGCTCTGTGGCCGACCGCTGCAGG 540
DB 481 GCCTCACAACCACTCCCTGGAGCGCTACCTGCACATGCTCTGTGGCCGACCGCTGCAGG 540

QY 541 CTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCCAGTATGCTT 600
DB 541 CTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCCAGTATGCTT 600

QY 601 CCTACCATGGCAGAGTCTGAATCCATCTCTTTTGGATTAACTCAACACAGGAT 660
DB 601 CCTACCATGGCAGAGTCTGAATCCATCTCTTTTGGATTAACTCAACACAGGAT 660

QY 661 TTGAACGGGCGAGTAAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGCGAG 720
DB 661 TTGAACGGGCGAGTAAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGCGAG 720

QY 721 AACGTGACCTCTTGTGTAAGGAGTCCACGCAAGGAGTGAAGCCCTGTTTCAGGGAGATC 780

DB 721 AACGTGACCTCTTGTGTAAGGAGTCCACGCAAGGAGTGAAGCCCTGTTTCAGGGAGATC 780

QY 781 ACAGCCTCTCTGCGGTCTCCATCTCTATCAAACTGAACAGGAGACCGACCTTGGCTG 840

DB 781 ACAGCCTCTCTGCGGTCTCCATCTCTATCAAACTGAACAGGAGACCGACCTTGGCTG 840

QY 841 TCGTGTCCAGGAATGTCTAGTGTGATGCAAAATGCAAAAGGAGCGGAAGAAAGAAAGA 900

DB 841 TCGTGTCCAGGAATGTCTAGTGTGATGCAAAATGCAAAAGGAGCGGAAGAAAGAAAGA 900

QY 901 AAGTGACCAACAATAATCTCATTTGATGATGAGGAAGATGAGAGCAACTCTGGGACGTGT 960

DB 901 AAGTGACCAACAATAATCTCATTTGATGATGAGGAAGATGAGAGCAACTCTGGGACGTGT 960

QY 961 TTAAGAACACCTCGGGCAGGGAGAGTCAAGAGCAAACTCCGACCGCTCTCTCTGCA 1020

DB 961 TTAAGAACACCTCGGGCAGGGAGAGTCAAGAGCAAACTCCGACCGCTCTCTCTGCA 1020

QY 1021 ATATCATGTCCGCTTTGAAAGCCCTTTCGGGCCCTTAACCTCAATGGAATCAAGAGCAAA 1080

DB 1021 ATATCATGTCCGCTTTGAAAGCCCTTTCGGGCCCTTAACCTCAATGGAATCAAGAGCAAA 1080

QY 1081 CTCATGGAATAATGATTCCTCTGTGAAACGGGAGTTTGGGTACCAAGACTTGTATGT 1140

DB 1081 CTCATGGAATAATGATTCCTCTGTGAAACGGGAGTTTGGGTACCAAGACTTGTATGT 1140

QY 1141 GAAAGCATCTGATGATGAAGATGTGATGAAACCAAGATGACGTGTATGGAATCTCATC 1200

DB 1141 GAAAGCATCTGATGATGAAGATGTGATGAAACCAAGATGACGTGTATGGAATCTCATC 1200

QY 1201 AGGAAGAACACAGGGGCCACTCGGAGTCCGCCGAGAAAGCCACTGGAGGGAACACCTG 1260

DB 1201 AGGAAGAACACAGGGGCCACTCGGAGTCCGCCGAGAAAGCCACTGGAGGGAACACCTG 1260

QY 1261 CCTCTCCAGATGCAAGCTGGGCTCCGCTGAGGTGCTGCAAACTCTCGACATCTCT 1320

DB 1261 CCTCTCCAGATGCAAGCTGGGCTCCGCTGAGGTGCTGCAAACTCTCGACATCTCT 1320

QY 1321 CTTTCCCTGTCAAGTGGGCTCTTACAGCCAGCAGATGCCCTCGGAAGCCCTGGA 1380

DB 1321 CTTTCCCTGTCAAGTGGGCTCTTACAGCCAGCAGATGCCCTCGGAAGCCCTGGA 1380

QY 1381 GAAAGGAGACAGGACAGGACCAAGTCTTCCCGGATCTTGGGATCTCGGTACAGTGTGA 1440

DB 1381 GAAAGGAGACAGGACAGGACCAAGTCTTCCCGGATCTTGGGATCTCGGTACAGTGTGA 1440

QY 1441 AGCCAGCTCTCCAGGCCAGGAAAGTCTCTGAGCAGCTGTACTTCTGCTCAGTGTCCA 1500

DB 1441 AGCCAGCTCTCCAGGCCAGGAAAGTCTCTGAGCAGCTGTACTTCTGCTCAGTGTCCA 1500

QY 1501 GAGTCCATGACAAATTAGTGAATCGGCCAGGCCACTGTGCGCATGATGAACAGGAGGAT 1560

DB 1501 GAGTCCATGACAAATTAGTGAATCGGCCAGGCCACTGTGCGCATGATGAACAGGAGGAT 1560

QY 1561 GAGTGTGAGGAGGAGACAGATCACTCGGAAACCTGCTCCAGCTGAGATGAGCACTCA 1620

DB 1561 GAGTGTGAGGAGGAGACAGATCACTCGGAAACCTGCTCCAGCTGAGATGAGCACTCA 1620

QY 1621 GCCCGCTCTCCGCGCAAGAGGTGACACCTTGAAGAGGAGTGGTGAACAGGAGGAGCGG 1680

DB 1621 GCCCGCTCTCCGCGCAAGAGGTGACACCTTGAAGAGGAGTGGTGAACAGGAGGAGCGG 1680

QY 1681 CAGGCAATGAAGGTCCAGGCGCTGCGGAGCTATCTTTGCTATTTTGTGAGGAGATCTAA 1740

DB 1681 CAGGCAATGAAGGTCCAGGCGCTGCGGAGCTATCTTTGCTATTTTGTGAGGAGATCTAA 1740

QY 1741 CCCACAGTGAAGAACCATGTGTGAGAAATGGAAGGAGAGAGAAATCAACAGTTCCTGA 1800

DB 1741 CCCACAGTGAAGAACCATGTGTGTGAGAAATGGAAGGAGAGAGAAATCAACAGTTCCTGA 1800

QY 1801 TAGTCTCTATTGAGCTCTCTGATTCAGTCTTTCTGAGCTGTCTTCTCTGAGCTTTT 1860

Db 1801 TAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCTGAAGCTGTGTTTCTCTGACACTTTT 1860
Qy 1861 CATGTATGTGAGCCAAATAAATGCTTTTCATTCCTTTGAAAAA 1906
Db 1861 CATGTATGTGAGCCAAATAAATGCTTTTCATTCCTTTGAAAAA 1906

RESULT 3
US-10-070-255-4
; Sequence 4, Application US/10070255
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolay
; APPLICANT: SINHA, Indranil
; APPLICANT: LEU, Stefan
; TITLE OF INVENTION: IREN PROTEIN, ITS PREPARATION AND USE
; FILE REFERENCE: WALLACH=28
; CURRENT APPLICATION NUMBER: US/10/070,255
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/IL00/00517
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: IL 131719
; PRIOR FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-070-255-4

Query Match 91.4%; Score 1742; DB 47; Length 1782;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1775; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

Qy 118 ATGAGCGGATCNCNGACNATGCAAAAGACAATTTCTGCTGAGCGACTGCTGGATGCA 177
Db 1 ATGAGCGGATCAGAGAACAAATGCAAAAGACAATTTCTGCTGAGCGACTGCTGGATGCA 60

Qy 178 GTGAAACAGTGCAGATCCGCTTTNGAGGGAGAGAGATTCCTCGGATTCGGACAGC 237
Db 61 GTGAAACAGTGCAGATCCGCTTTNGAGGGAGAGAGATTCCTCGGATTCGGACAGC 120

Qy 238 AGGGTCACTGTCTGTGTGCCAGTTTGAAGCCGCTCTGAGCATGCTTTGAAGAGAGT 297
Db 121 AGGGTCACTGTCTGTGTGCCAGTTTGAAGCCGCTCTGAGCATGCTTTGAAGAGAGT 180

Qy 298 CGAGGATTGGCACTCAGAGCGGCGAGATCAAGCAGGCGCGGCTTTGCCAGCAAAACC 357
Db 181 CGAGGATTGGCACTCAGAGCGGCGAGATCAAGCAGGCGCGGCTTTGCCAGCAAAACC 240

Qy 358 GAAACAGAGCCGCTGTCTGTGTAAGTCTGTAAGGAGTCTCAACAGACACAGCTGCAG 417
Db 241 GAAACAGAGCCGCTGTCTGTGTAAGTCTGTAAGGAGTCTCAACAGACACAGCTGCAG 300

Qy 418 CGCTTCTACTCCCTGCGCCACATCGCCTCAGAGCTGGCGCGGGTCCGCTGGCTGCGC 477
Db 301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGAGCTGGCGCGGGTCCGCTGGCTGCGC 360

Qy 478 TGTGCCCTCAACGAACATCTCTGGAGCGGTACTGTCAGCATGCTCTGGCCGACCGCTGC 537
Db 361 TGTGCCCTCAACGAACATCTCTGGAGCGGTACTGTCAGCATGCTCTGGCCGACCGCTGC 420

Qy 538 AGGCTGAGCACTTTTATGAAGTCTGTCTTTGATGATGATGATGATGATGATGATGATG 597
Db 421 AGGCTGAGCACTTTTATGAAGTCTGTCTTTGATGATGATGATGATGATGATGATGATG 480

Qy 598 CTTTCTACCATGCGCAGCGTCTGAATCCATCTCTTTGCGATTAAACATCGACAAACAG 657
Db 481 CTTTCTACCATGCGCAGCGTCTGAATCCATCTCTTTGCGATTAAACATCGACAAACAG 540

Qy 658 GATTGAAACGGGAGAGTAAAGTTTGTCTCCACCGTTTTCAGACTCTTAAAGGAGTCAACG 717
Db 1

Db 541 GATTGAAACGGGAGAGTAAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACG 600
Qy 718 CAGAAACGTGACCTCTCTGCTGAAAGGAGTCCACGCAAGGAGTGCAGACGCTGTTTCAGGGAG 777
Db 601 CAGAAACGTGACCTCTCTGCTGAAAGGAGTCCACGCAAGGAGTGCAGACGCTGTTTCAGGGAG 660

Qy 778 ATCACAGCCTCTCTGCTGCTCTCATCTCATCAAACTGAAACAGGAGACCGA-CCCTTG 836
Db 661 ATCACAGCCTCTCTGCTGCTCTCATCTCATCAAACTGAAACAGGAGACCGACCCCTTG 720

Qy 837 CTTGTGCTGTCCAGGAATGTCAGTGTGATGCCAAATGCAAAAGGAGCGGAAAGAAA 896
Db 721 CTTGTGCTGTCCAGGAATGTCAGTGTGATGCCAAATGCAAAAGGAGCGGAAAGAAA 780

Qy 897 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTTGGGGAC 956
Db 781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTTGGGGAC 840

Qy 957 GTGTTTAAAGAACACCTTGGGGAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCT 1016
Db 841 GTGTTTAAAGAACACCTTGGGGAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCT 900

Qy 1017 GTCAATATCATGTCCGCTTTGAAAGCCCTTGGGAGCCCTTAACTCCAAATGGAAGTCAAGC 1075
Db 901 GTCAATATCATGTCCGCTTTGAAAGCCCTTGGGAGCCCTTAACTCCAAATGGAAGTCAAGC 960

Qy 1076 AGCAACTCATGGAATAATGATTCCTCTGTTTGAACGGGGAGTTTGGGTACCAAGAGCTT 1135
Db 961 AGCAACTCATGGAATAATGATTCCTCTGTTTGAACGGGGAGTTTGGGTACCAAGAGCTT 1020

Qy 1136 GATGTGAAAGCATCGATGATGAAGATGAGTGAATAAAGCAAGATGACGTGTATGGAAC 1195
Db 1021 GATGTGAAAGCATCGATGATGAAGATGAGTGAATAAAGCAAGATGACGTGTATGGAAC 1080

Qy 1196 TCATCAGGAGGAGACACAGGGGCGCTCGGAGTCCGCCAGAGACCACTGGAAGGAAAC 1255
Db 1081 TCATCAGGAGGAGACACAGGGGCGCTCGGAGTCCGCCAGAGACCACTGGAAGGAAAC 1140

Qy 1256 ACCTGCTCTCCAGATGACACAGCTGGGCTCCGCTGAAGGTGCTGCAATGACTCTCGAC 1315
Db 1141 ACCTGCTCTCCAGATGACACAGCTGGGCTCCGCTGAAGGTGCTGCAATGACTCTCGAC 1200

Qy 1316 ATCTCTCTCTGCTAGTGGCTGCTCTACAGCCAGCAGATGCCCTCGGAAGC 1375
Db 1201 ATCTCTCTCTGCTAGTGGCTGCTCTACAGCCAGCAGATGCCCTCGGAAGC 1260

Qy 1376 CTGGAAACGGGACAGGACAGAGGACAGCTTCTCCGGATCCTGGACTTCGGTACAGT 1435
Db 1261 CTGGAAACGGGACAGGACAGAGGACAGCTTCTCCGGATCCTGGACTTCGGTACAGT 1320

Qy 1436 GTGGAAGCCAGCTCTCCAGCCACGGAAGTCTCTGAGCAGCCTGTTA-CTTCTGCTCA 1494
Db 1321 GTGGAAGCCAGCTCTCCAGCCACGGAAGTCTCTGAGCAGCCTGTTA-CTTCTGCTCA 1380

Qy 1495 GTGCCAGAGTCCATGACAAATTAGTGAATGCGCAGGCCACTGTGGCCATGATGAACAGG 1554
Db 1381 GTGCCAGAGTCCATGACAAATTAGTGAATGCGCAGGCCACTGTGGCCATGATGAACAGG 1440

Qy 1555 AAGGATGAGCTGGAGGAGAGAACAGATCACTGCGAAACCTGTGACGGTGAAGTGGAG 1614
Db 1441 AAGGATGAGCTGGAGGAGAGAACAGATCACTGCGAAACCTGTGACGGTGAAGTGGAG 1500

Qy 1615 CACTCAGCGCTCCGCGCAAGAGTGGACACCTTGAAGAGGAGTGGCTGAACAGGAG 1674
Db 1501 CACTCAGCGCTCCGCGCAAGAGTGGACACCTTGAAGAGGAGTGGCTGAACAGGAG 1560

Qy 1675 GAGCGCAGGCGATGAAGGTCCAGCGCTGGCCAGCTATCTTTGTATTTTGTGAGGAGA 1734
Db 1561 GAGCGCAGGCGATGAAGGTCCAGCGCTGGCCAGCTATCTTTGTATTTTGTGAGGAGA 1620

Qy 1735 TTCTAAACCCACGTGAGAACCATGTGGTGGAGAAATGGAGGGAGAGAGAAATCCACAGT 1794
Db 1621 TTCTAAACCCACGTGAGAACCATGTGGTGGAGAAATGGAGGGAGAGAGAAATCCACAGT 1680

QY 1795 TCCTGATAGTCTCATTTGAGCTCCCTGGATCCAGTCTTTCCTGAAGCTGTCTCTCG 1854
DB 1681 TCCTGATAGTCTCATTTGAGCTCCCTGGATCCAGTCTTTCCTGAAGCTGTCTCTCG 1740
QY 1855 ACTTTTCATGATGTGAGCCATAAATGCTTTTCATTCCTTG 1896
DB 1741 ACTTTTCATGATGTGAGCCATAAATGCTTTTCATTCCTTG 1782

RESULT 4
US-10-070-255-6
; Sequence 6, Application US/10070255
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolay
; APPLICANT: SINHA, Indranil
; APPLICANT: LEU, Stefan
; FILE OF INVENTION: IREN PROTEIN, ITS PREPARATION AND USE
; TITLE REFERENCE: WALLACH=28
; CURRENT APPLICATION NUMBER: US/10/070.255
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/IL00/00517
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: IL 131719
; PRIOR FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 2873
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-070-255-6

Query Match 81.6%; Score 1555; DB 47; Length 2873;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

QY 118 ATGAGCGGATCNCNGAACNATGACAAAGACAATTTCTGCTGGAGCACTGCTGGATGCA 177
DB 1 ATGAGCGGATCACAGAAACAATGACAAAGACAATTTCTGCTGGAGCACTGCTGGATGCA 60
QY 178 GTGAAACAGTCCAGATCCGCTTTNGAGGGAGAAAGAGATTGCTCGGATTCGACAGC 237
DB 61 GTGAAACAGTCCAGATCCGCTTTGAGGGAGAAAGAGATTGCTCGGATTCGACAGC 120
QY 238 AGGTCACCTGTGTGTGCCAGTTTGAAGCCGTCTGACAGCATGCTTGAAGAGAGT 297
DB 121 AGGTCACCTGTGTGTGCCAGTTTGAAGCCGTCTGACAGCATGCTTGAAGAGAGT 180
QY 298 CGAGGATGGCACTCAGAGCGGACGATCAAGCAGGCGGGCTTTGCCAGCAAAACC 357
DB 181 CGAGGATGGCACTCAGAGCGGACGATCAAGCAGGCGGGCTTTGCCAGCAAAACC 240
QY 358 GAAACAGAGCCGCTGTCTGCTGTAAGAGAGTCTTCAACAGACGAGCTGCAG 417
DB 241 GAAACAGAGCCGCTGTCTGCTGTAAGAGAGTCTTCAACAGACGAGCTGCAG 300
QY 418 CGCTTCTACTCCCTGCGGCAATCGCTCAGACGTTGGGCGGGGTGCGGCTGCTGCGC 477
DB 301 CGCTTCTACTCCCTGCGGCAATCGCTCAGACGTTGGGCGGGGTGCGGCTGCTGCGC 360
QY 478 TGTCCCTCAACGAACACTCCCTGGAGCGGTACTGTCACATGCTCTGCGGCGGCGCTGC 537
DB 361 TGTCCCTCAACGAACACTCCCTGGAGCGGTACTGTCACATGCTCTGCGGCGGCGCTGC 420
QY 538 AGGCTGAGCACTTTTATGAAGATGCTTTTGTGATGATGAAGAAAGTCCAGTATG 597
DB 421 AGGCTGAGCACTTTTATGAAGATGCTTTTGTGATGATGAAGAAAGTCCAGTATG 480
QY 598 CTTCTACCATGGCAGAGGTCTGAACCTCATCTCTTTGGGATTAATCATCGACAAAG 657
DB 481 CTTCTACCATGGCAGAGGTCTGAACCTCATCTCTTTGGGATTAATCATCGACAAAG 540

QY 658 GATTTGAACGGGACAGATGAAGTTTGTCTCCCAACCGTTTTCAGACCTCTTAAAGGAGTCAACG 717
DB 541 GATTTGAACGGGACAGATGAAGTTTGTCTCCCAACCGTTTTCAGACCTCTTAAAGGAGTCAACG 600
QY 718 CAGAACGTGACCTCTCTTGTGAAGAGTCCACAGCAAGAGTGACAGCCTCTTCAGGGAG 777
DB 601 CAGAACGTGACCTCTCTTGTGAAGAGTCCACAGCAAGAGTGACAGCCTCTTCAGGGAG 660
QY 778 ATCAGACGCTCTCTGCGGTCTCCATCTCATCAAACTTGAACAGGAGACCGA-CCCTTG 836
DB 661 ATCAGACGCTCTCTGCGGTCTCCATCTCATCAAACTTGAACAGGAGACCGACCCCTTG 720
QY 837 CTTCTGCTGTCCAGGAATGTCACTGCTGATGCCAAATGCAAAAGAGAGCGGAAAGAAA 896
DB 721 CTTCTGCTGTCCAGGAATGTCACTGCTGATGCCAAATGCAAAAGAGAGCGGAAAGAAA 780
QY 897 AAGAAAGTGACCAACATAATCTCATTTTGTGATGATGAGGAAGATGACAGAACTCTGGGAC 956
DB 781 AAGAAAGTGACCAACATAATCTCATTTTGTGATGATGAGGAAGATGACAGAACTCTGGGAC 840
QY 957 GTGTTTAAAAAGACACCTGGGCGAGGGAGAGCTTCAGAGGACAACTCCGACCGCTCTCT 1016
DB 841 GTGTTTAAAAAGACACCTGGGCGAGGGAGAGCTTCAGAGGACAACTCCGACCGCTCTCT 900
QY 1017 GTCAATATCATGTCCGCTTTGAAAGCCCTTCGGGCTTAATCCCAATGGAA-TCAGAGC 1075
DB 901 GTCAATATCATGTCCGCTTTGAAAGCCCTTCGGGCTTAATCCCAATGGAAAGTCAGAGC 960
QY 1076 AGCAACTCATGGAATAATGATTCCTGTCTTTGAAAGGGGAGTTTGGGTACCAAGACTT 1135
DB 961 AGCAACTCATGGAATAATGATTCCTGTCTTTGAAAGGGGAGTTTGGGTACCAAGAGCTT 1020
QY 1136 GATGTGAAAAAGCATCGATGATGAAGATGTGCGATGAAAGCAAGATGACGTGTATGGAAC 1195
DB 1021 GATGTGAAAAAGCATCGATGATGAAGATGTGCGATGAAAGCAAGATGACGTGTATGGAAC 1080
QY 1196 TCATCAGAGGAGGACACAGGGGCCACTCGAGTCCGCCGAGGACCACTTGGAGGGAAC 1255
DB 1081 TCATCAGAGGAGGAGACACAGGGGCCACTCGAGTCCGCCGAGGACCACTTGGAGGGAAC 1140
QY 1256 ACCTGCTCTCCAGATGACAGCTGGGCTCCGCTGAAGGTGTGCACAATAGCTCCGAC 1315
DB 1141 ACCTGCTCTCCAGATGACAGCTGGGCTCCGCTGAAGGTGTGCACAATAGCTCCGAC 1200
QY 1316 ATCTCTTCTCTGTGCTGCTGGGCTCTCTACAGCCAGCAGATGCCCCCTCGGAAGC 1375
DB 1201 ATCTCTTCTCTGTGCTGCTGGGCTCTCTACAGCCAGCAGATGCCCCCTCGGAAGC 1260
QY 1376 CTGGAGAACGGGACAGGACCAAGGACCAAGTCTCCCGATCCTGGAATTCGGTACAGT 1435
DB 1261 CTGGAGAACGGGACAGGACCAAGGACCAAGTCTCCCGATCCTGGAATTCGGTACAGT 1320
QY 1436 GTGGAAAGCCAGCTCTCAGGCCACGGAAGTCTCTGAGCAGCCTGTTA-CTTCTGCCCTCA 1494
DB 1321 GTGGAAAGCCAGCTCTCAGGCCACGGAAGTCTCTGAGCAGCCTGTTAATCTTCTGCTCA 1380
QY 1495 GTGCCAGAGTCCATGACAAATAGTGAACCTGCGGCAAGGCACTGTGGCCATGATGAACAG 1554
DB 1381 GTGCCAGAGTCCATGACAAATAGTGAACCTGCGGCAAGGCACTGTGGCCATGATGAACAG 1440
QY 1555 AAGGATGAGCTGAGGAGGAGAACAGATCACTCGGAACCTGCTCGACGCTGAGATGGAG 1614
DB 1441 AAGGATGAGCTGAGGAGGAGAACAGATCACTCGGAACCTGCTCGACGCTGAGATGGAG 1500
QY 1615 CACTCAGCCCGCTCCGGCAAGAGGTGGACACCTTGAAGAGAGAGGTGGTGGTGAACAGAG 1674
DB 1501 CACTCAGCCCGCTCCGGCAAGAGGTGGACACCTTGAAGAGAGAGGTGGTGGTGAACAGAG 1560
QY 1675 GAGCGGAGGGCATGAAAGTCCAGGGCGCTGGCCAG 1709
DB 1561 GAGCGGAGGGCATGAAAGTCCAGGGCGCTGGCCAG 1595

RESULT 5
US-10-070-255-5
; Sequence 5, Application US/10070255
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolay
; APPLICANT: SINHA, Indranil
; APPLICANT: LEU, Stefan
; TITLE OF INVENTION: IREN PROTEIN, ITS PREPARATION AND USE
; FILE REFERENCE: WALLACH=28
; CURRENT APPLICATION NUMBER: US/10/070,255
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/IL00/00517
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: IL 131719
; PRIOR FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 3139
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-070-255-5

Query Match 81.6%; Score 1555; DB 47; Length 3139;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

QY	118	ATCAGCGGATCNCNGACNATGACAAAGACATTTCTGCTGAGCGAGTCTGGATGCA	177
DB	1	ATCAGCGGATCAGACAAATGACAAAGACATTTCTGCTGAGCGAGTCTGGATGCA	60
QY	178	GTGAAACAGTCCAGATCCCTTTNGAGGAGAAAGAGATTCCTCGGATTCGACAGC	237
DB	61	GTGAAACAGTCCAGATCCCTTTNGAGGAGAAAGAGATTCCTCGGATTCGACAGC	120
QY	238	AGGCTCACCTGTCTGTGCCCCAGTTTGAAGCCGTCTCGACGATGGCTTGAAGAGAGT	297
DB	121	AGGCTCACCTGTCTGTGCCCCAGTTTGAAGCCGTCTCGACGATGGCTTGAAGAGAGT	180
QY	298	CGAGGATTGGCACTCAGCGGGCAGCGATCAACAGCAGCAGCGGGCTTTGCGACAAACC	357
DB	181	CGAGGATTGGCACTCAGCGGGCAGCGATCAACAGCAGCAGCGGGCTTTGCGACAAACC	240
QY	358	GAAACAGAGCCGTGTTCTGTGTAAGTACTAGTGAAGAGTCTCTCAACAGCAGCTGCAG	417
DB	241	GAAACAGAGCCGTGTTCTGTGTAAGTACTAGTGAAGAGTCTCTCAACAGCAGCTGCAG	300
QY	418	CGCTTCTACTCTCGGCCACATCGCTCAGACGTGGCGGGGTGGCGCTGGCTGGCG	477
DB	301	CGCTTCTACTCTCGGCCACATCGCTCAGACGTGGCGGGGTGGCGCTGGCTGGCG	360
QY	478	TGTGCCCTCAACGAACTCCCTGGAGCGGTACTGACATGCTCTGCGCGCAGCCGTGC	537
DB	361	TGTGCCCTCAACGAACTCCCTGGAGCGGTACTGACATGCTCTGCGCGCAGCCGTGC	420
QY	538	AGGCTGAGCACTTTTATGAAGTCTGTTTGTGATGATGAAGAGTCCAGTATG	597
DB	421	AGGCTGAGCACTTTTATGAAGTCTGTTTGTGATGATGAAGAGTCCAGTATG	480
QY	598	CTTCTACCATGGCAGCGGTCTGAATCCATCTCTTTGCGATTAAACATCGACAAAG	657
DB	481	CTTCTACCATGGCAGCGGTCTGAATCCATCTCTTTGCGATTAAACATCGACAAAG	540
QY	658	GATTGAAACGGGAGAGTAAAGTTTGTCCACCGTTTCAGACCTCTTAAAGAGTCAACG	717
DB	541	GATTGAAACGGGAGAGTAAAGTTTGTCCACCGTTTCAGACCTCTTAAAGAGTCAACG	600
QY	718	CAGAACTGACCTCTTGTGTAAGAGTCCACCGAAGGAGTGAAGCCTGTTTCAGGGAG	777
DB	601	CAGAACTGACCTCTTGTGTAAGAGTCCACCGAAGGAGTGAAGCCTGTTTCAGGGAG	660

QY	778	ATCAGAGCCTCTCTGCGCTCTCCATCTCATCAAACTGAAACAGAGACCGA-CCTTTG	836
DB	661	ATCAGAGCCTCTCTGCGCTCTCCATCTCATCAAACTGAAACAGAGACCGA-CCTTTG	720
QY	837	CCTGTCGTCTCAGGAATGTCAAGTGTGATGCCAAATGCCAAAGAGCGGAAGAGAA	896
DB	721	CCTGTCGTCTCAGGAATGTCAAGTGTGATGCCAAATGCCAAAGAGCGGAAGAGAA	780
QY	897	AAGAAAGTGAACCAACATAATCTCATTTGATGATGAGAAAGATGAGCAACTCTGGGAC	956
DB	781	AAGAAAGTGAACCAACATAATCTCATTTGATGATGAGAAAGATGAGCAACTCTGGGAC	840
QY	957	GTGTTTAAAAAGACACCTGGGAGAGAGCTCAGAGACAACTCCGACCGCTCTCT	1016
DB	841	GTGTTTAAAAAGACACCTGGGAGAGAGCTCAGAGACAACTCCGACCGCTCTCT	900
QY	1017	GTCAATATCATGTCCGCTTTTGAAGCCCTTTCGGGCTTAATCTCAATGGAATCGAGC	1075
DB	901	GTCAATATCATGTCCGCTTTTGAAGCCCTTTCGGGCTTAATCTCAATGGAATCGAGC	960
QY	1076	AGCAACTCATGGAATAATGATTCCTGTCTTTGAAACGGGAGTTTGGGTACCAAGAGCTT	1135
DB	961	AGCAACTCATGGAATAATGATTCCTGTCTTTGAAACGGGAGTTTGGGTACCAAGAGCTT	1020
QY	1136	GATGGAAGAGATCGATGATGAAGATGTGATGAAAGAGAGATGACGTGTATGGAAC	1195
DB	1021	GATGGAAGAGATCGATGATGAAGATGTGATGAAAGAGAGATGACGTGTATGGAAC	1080
QY	1196	TCATCAGGAAGAGACACAGGGGCCACTCGAGTCCGCCGAGAGCCACTTGGAGGGAAC	1255
DB	1081	TCATCAGGAAGAGACACAGGGGCCACTCGAGTCCGCCGAGAGCCACTTGGAGGGAAC	1140
QY	1256	ACCTGCTCTCCAGATGACAGTGGGCTCCGCTGAAGGTGTGCACATGATCTCCGAC	1315
DB	1141	ACCTGCTCTCCAGATGACAGTGGGCTCCGCTGAAGGTGTGCACATGATCTCCGAC	1200
QY	1316	ATCCTCTTCCCTGTCAAGTGGGCTCTTACAGCCAGAGATGCCCCCTCGGAGC	1375
DB	1201	ATCCTCTTCCCTGTCAAGTGGGCTCTTACAGCCAGAGATGCCCCCTCGGAGC	1260
QY	1376	CTGGAGACGGGACAGGACAGAGGACAGTTCCTCCGGATCCTCGACTTCGGTACAGT	1435
DB	1261	CTGGAGACGGGACAGGACAGAGGACAGTTCCTCCGGATCCTCGACTTCGGTACAGT	1320
QY	1436	GTGGAGACGGGACAGTCTCCAGGCCACGGAAGTCTCTGAGCAGCCTGTTA-CTTCTGCTCA	1494
DB	1321	GTGGAGACGGGACAGTCTCCAGGCCACGGAAGTCTCTGAGCAGCCTGTTA-CTTCTGCTCA	1380
QY	1495	GTGCCAGAGTCCATGACAAATTAGTGAATGCGCCAGGCCACTGTGGCCATGATGAACAGG	1554
DB	1381	GTGCCAGAGTCCATGACAAATTAGTGAATGCGCCAGGCCACTGTGGCCATGATGAACAGG	1440
QY	1555	AAGGATGAGTGGGAGGAGAGACAGATCACTGCGAAACCTGCTCGAGTGGATGGAG	1614
DB	1441	AAGGATGAGTGGGAGGAGAGACAGATCACTGCGAAACCTGCTCGAGTGGATGGAG	1500
QY	1615	CACTCAGCCGCTCCGCGAAGAGTGGACACTTGAAGAGGAGTGGCTGAACAGGAG	1674
DB	1501	CACTCAGCCGCTCCGCGAAGAGTGGACACTTGAAGAGGAGTGGCTGAACAGGAG	1560
QY	1675	GAGCGCAGGAGATGAAGTCCAGCGCTGGCCAG	1709
DB	1561	GAGCGCAGGAGATGAAGTCCAGCGCTGGCCAG	1595

RESULT 6
US-09-629-469A-14016
; Sequence 14016, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI

; SOFTWARE: pt_CT_genes Version 1.01

; SEQ ID NO 9141

; LENGTH: 1516

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1196)...(1516)

; OTHER INFORMATION: similar to gil1710216 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters

US-09-488-725B-9141

Query Match 57.3%; Score 1092.4; DB 22; Length 1516;

Best Local Similarity 97.6%; Pred. No. 7e-259;

Matches 1137; Conservative 0; Mismatches 23; Indels 5; Gaps 3;

```
QY 77 GCCGGCGCGCGCGCCGANGCACCAGCCCGCGGGANAGCNCATGAGCGGATCNCNGAACN 136
DB 27 GCGCGCGCGCGCGCGAGCACCAGCCCGCGGGAGAGCAGCAGTGGAGGATCAGAGCA 86
QY 137 ATGACAAAAGACAAATTTCTGCTGGAGCGACTGCTGGATGAGTGAAGACAGTGCAGATCC 196
DB 87 ATGACAAAAGACAAATTTCTGCTGGAGCGACTGCTGGATGAGTGAAGACAGTGCAGATCC 146
QY 197 GCTTTNAGGAGGAGAAAGGAGATTCCTCGGATTCGAGATTCGAGATTCGAGATTCGAG 256
DB 147 GCTTTNAGGAGGAGAAAGGAGATTCCTCGGATTCGAGATTCGAGATTCGAGATTCGAG 206
QY 856 TCAGTGTCTGATGTCAAATGCAAAAGGAGCGGAGAGAAAGAAAGATGACCAACATTA 915
DB 804 TCAGTGTCTGATGTCAAATGCAAAAGGAGCGGAGAGAAAGAAAGATGACCAACATTA 863
QY 916 TCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTTAAAAAGACACCTG 975
```

```
DB 864 ACTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGACATGTTTAAAAAGACACCTG 923
QY 976 GCGCAGGGGAGAGCTCAGAGGACAACTCCAGCGCTCTCTGTCAATATCATGTCCGCCT 1035
DB 924 GCGCAGGGGAGAGCTCAGAGGACAACTCCAGCGCTCTCTGTCAATATCATGTCCGCCT 983
QY 1036 TTGAAAGCCCTTTCGGGCTTAATCTCAATGAA-TCAGAGCAGCAACTCATGGAATTCG 1094
DB 984 TTGAAAGCCCTTTCGGGCTTAATCTCAATGAA-TCAGAGCAGCAACTCATGGAATTCG 1043
QY 1095 ATTCCCTGTCTTTGAAACGGGAGCTTTGGGTACAGAGCTTGTGTAAGAGCATCGATG 1154
DB 1044 ATTCCCTGTCTTTGAAACGGGAGCTTTGGGTACAGAGCTTGTGTAAGAGCATCGATG 1103
QY 1155 ATGAAGATGTGGATGAAACGAAAGATGACCTGTATGAAACTCATCAGGAAGAGGACACA 1214
DB 1104 ATGAAGATGTGGATGAAACGAAAGATGACCTGTATGAAACTCATCAGGAAGAGGACACA 1163
QY 1215 GGGGCCACTCGGAGTCGCCCGAGAA 1239
DB 1164 GGGGCCACTCAGAGTCGCCCGAGAA 1188
```

RESULT 10

US-09-552-317-9141

; Sequence 9141, Application US/09552317

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Tillinghast, John

; APPLICANT: Sinku, Ankura

; APPLICANT: Liu, Chenghua

; APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: Novel Nucleic Acids

; TITLE OF INVENTION: and Polypeptides

; FILE REFERENCE: 784CIP

; CURRENT APPLICATION NUMBER: US/09/552,317

; CURRENT FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 10289

; SOFTWARE: pt_CT_genes Version 1.01

; SEQ ID NO 9141

; LENGTH: 1516

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1196)...(1516)

; OTHER INFORMATION: similar to gil1710216 in the genepept database release 114,

; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters

US-09-552-317-9141

Query Match 57.3%; Score 1092.4; DB 26; Length 1516;

Best Local Similarity 97.6%; Pred. No. 7e-259;

Matches 1137; Conservative 0; Mismatches 23; Indels 5; Gaps 3;

```
QY 77 GCGCGCGCGCGCGCANGCACCAGCCCGCGGGANAGCNCATGAGCGGATCNCNGAACN 136
DB 27 GCGCGCGCGCGCGCGAGCACCAGCCCGCGGGAGAGCAGCAGTGGAGGATCAGAGCA 86
QY 137 ATGACAAAAGACAAATTTCTGCTGGAGCGACTGCTGGATGAGTGAAGACAGTGCAGATCC 196
DB 87 ATGACAAAAGACAAATTTCTGCTGGAGCGACTGCTGGATGAGTGAAGACAGTGCAGATCC 146
QY 197 GCTTTNAGGAGGAGAAAGGAGATTCCTCGGATTCGAGATTCGAGATTCGAGATTCGAG 256
DB 147 GCTTTNAGGAGGAGAAAGGAGATTCCTCGGATTCGAGATTCGAGATTCGAGATTCGAG 206
QY 257 CCCAGTTTGAAGCCGCTTCGAGCATGGCTTGAAGAGGAGTCCAGGATTCGAGATTCGAG 316
DB 207 CCCAGTTTGAAGCCGCTTCGAGCATGGCTTGAAGAGGAGTCCAGGATTCGAGATTCGAG 266
```

317 CGGAGCGATCAAGCAGGAGCGGGCTTTGGCAGCAAAACGAAACAGAGCGCGGTCT 376
326 CGGAGCGATCAAGCAGGAGCGGGCTTTGGCAGCAAAACGAAACAGAGCGCGGTCT 326
377 GGTACTACGTGAAGGAGTCTCTCAACAGCAGAGCTGACGGCTTCTACTCCCTGGCC 436
327 GGTACTACGTGAAGGAGTCTCTCAACAGCAGAGCTGACGGCTTCTACTCCCTGGCC 386
437 ACATCGCTCAGAGTGGGGCGGGTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 496
387 ACATCGCTCAGAGTGGGGCGGGTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 446
497 CCTCGAGCGTACCTGACATGCTCTCTGGCGGAGCTGAGGCTGAGCACTTTTATG 556
447 CCTGGAGCGTACCTGACATGCTCTCTGGCGGAGCTGAGGCTGAGCACTTTTATG 506
557 AAGACTGCTTTTGTGATGATGAAGAAAGTCCAGTATGCTTCTTACCATGGAGGAG 616
507 AAGACTGCTTTTGTGATGATGAAGAAAGTCCAGTATGCTTCTTACCATGGAGGAG 566
617 GTCTGAATCCATCTCTTGGATTAACATCGACAAAGGATTTGAACGGGAGAGTA 676
567 GTCTGAATCCATCTCTTGGATTAACATCGACAAAGGATTTGAACGGGAGAGTA 626
677 AGTTTGTCTCCACCGCTTTCAGACCTCTTAAAGGAGTCAACGAGAACTGACCTCTTGC 736
627 AGTTTGTCTCCACCGCTTTCAGACCTCTTAAAGGAGTCAACGAGAACTGTA--CCTTGC 683
737 TGAAGGAGTCCAAGGAGGAGTGAAGGAGTCTTTCAGGAGATCAACGCTTCTTGGCG 796
684 TGAAGGAGTCCAAGGAGGAGTGAAGGAGTCTTTCAGGAGATCAACGCTTCTTGGCA 743
797 TCTCCATCTCATCAACCTGAGAGGAGCGGA--CCCTTGCCTGCTGCTGCGAGGATG 855
744 TCTCCATCTCATCAACCTGAGAGGAGCGGA--CCCTTGCCTGCTGCTGCGAGGATG 803
856 TCAGTGTCTGATGCCAAATGCAAAAGGAGCGGAAGAAAGAAAGTGAACCAATTA 915
804 TCAGTGTCTGATGCCAAATGCAAAAGGAGCGGAAGAAAGTGAACCAATTA 863
916 TCTCATTTGATGATGAGGAGGAGTGAAGGAGTCTTGGGAGCGTGTAAAAAGACACCTG 975
864 ACTCATTTGATGATGAGGAGGAGTGAAGGAGTCTTGGGAGCGTGTAAAAAGACACCTG 923
976 GGGCAGGAGGAGTCAAGGAGCAACTCCGACCGCTCTCTGCTCAATATCATGTCCGCT 1035
924 GGGCAGGAGGAGTCAAGGAGCAACTCCGACCGCTCTCTGCTCAATATCATGTCCGCT 983
1036 TTGAAAGCCCTTCGGGCTTAATCCAATGGAA--TCAGAGCAGCAACTCATGGAAATG 1094
984 TTGAAAGCCCTTCGGGCTTAATCCAATGGAA--TCAGAGCAGCAACTCATGGAAATG 1043
1095 ATTCCCTGCTTTGAAAGCGGAGTGGTACCAAGGAGTTCATGTAAGGAGCATGATG 1154
1044 ATTCCCTGCTTTGAAAGCGGAGTGGTACCAAGGAGTTCATGTAAGGAGCATGATG 1103
1155 ATGAGATGCTGATGAAGGAGGAGTGAAGTATGTAAGGAACTCATCAGGAGGAGGAGCA 1214
1104 ATGAGATGCTGATGAAGGAGGAGTGAAGTATGTAAGGAACTCATCAGGAGGAGGAGCA 1163
1215 GGGGCGACTCGGAGTCCGCCGAGAA 1239
1164 GGGGCGACTCAGAGTCCGCCGAGAA 1188

RESULT 11
US-09-758-472-1043
; Sequence 1043, Application US/09758472
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PH001
; CURRENT APPLICATION NUMBER: US/09/758,472

; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 9632
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1043
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (121)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (184)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (540)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1100)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1111)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-758-472-1043
Query Match 46.4%; Score 883.6; DB 33; Length 1137;
Best Local Similarity 98.4%; Pred. No. 2.9e-207;
Matches 949; Conservative 3; Mismatches 6; Indels 6; Gaps 6;
364 GAGCCCTGTTCTGGTACTACGTGAGGAGGTCTCTCAACAGCAGAGCTGCGCGCTTC 423
176 GAGCCCTGTTCTGGTACTACGTGAGGAGGTCTCTCAACAGCAGAGCTGCGAG-GCTTC 234
424 TACTCCCTGCGGCACATCGCTCAGAGCTGCGCGGCTCGCGCTCGCTGCGCTGCGC 483
235 TACTCCCTGCGGCACATCGCTCAGAGCTGCGCGGCTCGCGCTCGCTGCGCTGCGC 294
484 CTCAACGAACACATCCCTGGAGCGCTACCTGCACATGCTCTTGGCCGACCGCTGCGCTG 543
295 CTCAACGAACACATCCCTGGAGGCTACCTGCACATGCTCTTGGCCGACCGCTGCGCTG 354
544 AGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGGTCCAGTATGCTTCCT 603
355 AGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGGTCCAGTATGCTTCCT 414
604 ACCATGCGAGAGTCTGAATCCATCTCTTTCGGATTAAACATCGACAAACAGGATTG 663
415 ACCATGCGAGAGTCTGAATCCATCTCTTTCGGATTAAACATCGACAAACAGGATTG 474
664 AACGGCAGAGTAAGTTTGTCTCCACCGTTTTCAGACTCTTTAAAGAGTCAACCGAGAC 723
475 AACGGCAGAGTAAGTTTGTCTCCACCGTTTTCAGACTCTTTAAAGAGTCAACCGAGAA 534
724 GTGACCTCTTGTGAGGAGTCCACGAGAGTTCAGAGCTGAGCAGCTGTTTTCAGGAGATCACA 783
535 GTGACNTCTTGTGAGGAGTCCACGAGAGTTCAGAGCTGAGCAGCTGTTTTCAGGAGATCACA 594
784 GCCTCTCTGCGCTCTCCATCTCATCAAACTGAAACAGGAGACCGA--CCCTTGCCTGTC 842
595 GCCTCTCTGCGCTCTCCATCTCATCAAACTGAAACAGGAGACCGACCCCTTTCCTGTC 654
843 GTGTCAGGAAATGTCAAGTGTGATGCCAAATGCAAAAGGAGCGGAGGAGGAGGAGGAGGAG 902
655 GTGTCAGGAAATGTCAAGTGTGATGCCAAATGCAAAAGGAGCGGAGGAGGAGGAGGAGGAG 714
903 GTGACCAACATAATCTCATTTTGTGATGAGGAGATGAGCAGAACTCTTGGGGAGCTGTTT 962


```
Db 715 GTGACCAACATAATCTCTTTTGTATGAGGAAGATGACAGAACTCTGGGGACRTGTGT 774
Qy 963 AAAAAGACACCTTGGGGCAGGGGAGAGCTTCAGAGACAACTCCGACCGCTCTCTGTCAAT 1022
Db 775 AAAAAGACACCTTGGGGCAGGGGAGAGCTTCAGAGACAACTCCGACCGCTCTCTGTCAAT 834
Qy 1023 ATCATGTCGGCTTTGAAAGCCCTTCGGGCTTAATCTCAATGGAA-TCAGAGCAGCAAC 1081
Db 835 ATCATGTCGGCTTTGAAAGCCCTTCGGGCTTAATCTCAATGGAAAGTCAGAGCAGCAAC 894
Qy 1082 TCATGGAAAATTGATTCCTCTCTTTGAACGGGGAGTTTGGGTACCAAGCTTGTGTG 1141
Db 895 TCATGGAAAATTGATTCCTCTCTTTGAACGGGGAGTTTGGGTACCAAGCTTGTGTG 954
Qy 1142 AAAAGCATCGATGATGAAGATGTGGATGAAACGAAAGATGACGTGTATGGAAAATCATCA 1201
Db 955 AAAAGCATCGATGATGAAGATGKGGATGAAACGAAAGATGACGTGTATGGAAAATCATCA 1014
Qy 1202 GGAAGGAAGCAGGGGCACTCGGAGTCGCCGAGAGGCCACTGGAAAGGAAACACCTGC 1261
Db 1015 GGAAGGAAGCAGGGGCACTCGGAGTCG-CCGAGAGCCACTGGAAAGGAAACACCTGC 1073
Qy 1262 CTCTCCAGATGCACAGCTGGGCTCCGCTGAGGTGCTGACAAATGACTCCGACATCCTC 1321
Db 1074 CT-TTCCAGATGCACAGCTGGGCTCCGCTGAGGTGCTGACAAATGACTCCGACATCCTC 1131
Qy 1322 TTCC 1325
Db 1132 TTCC 1135
```

RESULT 12

```
US-10-235-926-1043
; Sequence 1043, Application US/10235926
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PH001CIN
; CURRENT APPLICATION NUMBER: US/10/235,926
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: 09/758,472
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 9632
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 1043
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (121)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (184)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (540)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1100)
; OTHER INFORMATION: n equals a,t,g, or c

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1111)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-235-926-1043

Query Match 46.4%; Score 883.6; DB 50; Length 1137;
Best Local Similarity 98.4%; Pred. No. 2.9e-207;
Matches 949; Conservative 3; Mismatches 6; Indels 6; Gaps 6;

Qy 364 GAGCCCGTGTTCGGTACTACGTGAAGSAGGTCTCTCAACAGCAGAGCTGCAGCGCTTC 423
Db 176 GAGCCCGTNTTCTGGTACTACGTGAAGSAGGTCTCTCAACAGCAGAGCTGCAG-GCTTC 234
Qy 424 TACTCCCTGGCCACATCGCTCTCAGACGTGGCGGGGTGCGGCTTGCGCTGTGCTGCC 483
Db 235 TACTCCCTGGCCACATCGCTCTCAGACGTGGCGGGGTGCGGCTTGCGCTGTGCTGCC 294
Qy 484 CTCAACGAACACTCTCTGGAGCGCTACTCTGACATGCTCTCTGGCCGACCGTGCAGGCTG 543
Db 295 CTCAACGAACACTCTCTGGAGCGCTACTCTGACATGCTCTCTGGCCGACCGTGCAGGCTG 354
Qy 544 AGCACTTTTATGAAGACTGCTCTTTTGTGATGATGAAGAAAGTCCAGTATCTTCCT 603
Db 355 AGCACTTTTATGAAGACTGCTCTTTTGTGATGATGAAGAAAGTCCAGTATCTTCCT 414
Qy 604 ACCATGGCAGCAGGTCTGAACCTCCATCTCTTTGCGATTAAACATCGAACCAAGGATTG 663
Db 415 ACCATGGCAGCAGGTCTGAACCTCCATCTCTTTGCGATTAAACATCGAACCAAGGATTG 474
Qy 664 AACGGGAGAGTAAGTTTGTCTCCACCGTTTTCAGACTCTTTAAAGAGTCAACGAGAAC 723
Db 475 AACGGGAGAGTAAGTTTGTCTCCACCGTTTTCAGACTCTTTAAAGAGTCAACGAGAAC 534
Qy 724 GTGACCTCTCTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTTCAGGGAGATCACA 783
Db 535 GTGACNTCTTGTGTAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTTCAGGGAGATCACA 594
Qy 784 GCCTCTCTGCGCTCTCCATCTCATCAACCTGAAACAGAGAGCCGA-CCCTTCCTGCTC 842
Db 595 GCCTCTCTGCGCTCTCCATCTCATCAACCTGAAACAGAGAGCCGACCCCTTTCCTGCTC 654
Qy 843 GTGTCCAGGAATGTCTAGTGTGATGCCAAATGCAAAAAGAGGCGGAAGAAAAAGAAA 902
Db 655 GTGTCCAGGAATGTCTAGTGTGATGCCAAATGCAAAAAGAGGCGGAAGAAAAAGAAA 714
Qy 903 GTGACCAACATAATCTCTTTGATGATGAGAGAGATCAGCAGACTCTGCGGAGCTGTGTT 962
Db 715 GTGACCAACATAATCTCTTTGATGATGAGAGAGATGAGCAGAACTCTTGGGGACRTGTGTT 774
Qy 963 AAAAAGACACCTGGGGCAGGGGAGAGCTTCAGAGACAACTCCGACCGCTCTCTGTCAAT 1022
Db 775 AAAAAGACACCTGGGGCAGGGGAGAGCTTCAGAGACAACTCCGACCGCTCTCTGTCAAT 834
Qy 1023 ATCATGTCGGCTTTGAAAGCCCTTCGGGCTTAATCTCAATGGAA-TCAGAGCAGCAAC 1081
Db 835 ATCATGTCGGCTTTGAAAGCCCTTCGGGCTTAATCTCAATGGAAAGTCAGAGCAGCAAC 894
Qy 1082 TCATGGAAAATTGATTCCTCTCTTTGAACGGGGAGTTTGGGTACCAAGCTTGTGTG 1141
Db 895 TCATGGAAAATTGATTCCTCTCTTTGAACGGGGAGTTTGGGTACCAAGCTTGTGTG 954
Qy 1142 AAAAGCATCGATGATGAAGATGTGGATGAAACGAAAGATGACGTGTATGGAAAATCATCA 1201
Db 955 AAAAGCATCGATGATGAAGATGKGGATGAAACGAAAGATGACGTGTATGGAAAATCATCA 1014
Qy 1202 GGAAGGAAGCAGGGGCACTCGGAGTCGCCGAGAGGCCACTGGAAAGGAAACACCTGC 1261
Db 1015 GGAAGGAAGCAGGGGCACTCGGAGTCG-CCGAGAGCCACTGGAAAGGAAACACCTGC 1073
Qy 1262 CTCTCCAGATGCACAGCTGGGCTCCGCTGAGGTGCTGACAAATGACTCCGACATCCTC 1321
Db 1074 CT-TTCCAGATGCACAGCTGGGCTCCGCTGAGGTGCTGACAAATGACTCCGACATCCTC 1131
```

```
QY 1322 TTCC 1325
Db 1132 TTCC 1135

RESULT 13
US-09-629-469A-18174
; Sequence 18174, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18174
; LENGTH: 6045
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (364)..(954)
US-09-629-469A-18174

Query Match 41.6%; Score 792.8; DB 28; Length 6045;
Best Local Similarity 96.8%; Pred. No. 1.5e-184;
Matches 852; Conservative 0; Mismatches 22; Indels 6; Gaps 4;

QY 363 AGAGCCCGTGTCTGGTACTAGCTGAAGGAGGTCTCTCAACAGCAGCAGCTGC-AGCGCT 421
Db 932 AAGCCCGTGTCTGGTACTAGCTGAAGGAGGTCTCTCAACAGCAGCAGCTGC-AGCGCT 991
QY 422 TTACTTCTCCGCCACATCGCTCAGAGTGGCGGGGGTGGCGCTGGCTGGCTGGT 481
Db 992 TTACTTCTCCGCCACATCGCTCAGAGTGGCGGGGGTGGCGCTGGCTGGCTGGT 1051
QY 482 CCTCAACCAACTCCCTGGAGCGCTACTCGACATGCTCTGGCCGCGCGCTGCAGGC 541
Db 1052 CCTCAACCAACTCCCTGGAGCGCTACTCGACATGCTCTGGCCGCGCGCTGCAGGC 1111
QY 542 TGAGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCCAGTATGCTTC 601
Db 1112 TCAGTACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCCAGTATGCTTC 1171
QY 602 CTACCATGGCAGCAGTCTGAACCTCCATCTCTTTGCGATTAAATCGACCAAGGATT 661
Db 1172 CTACCATGGCAGCAGTCTGAACCTCCATCTCTTTGCGATTAAATCGACCAAGGATT 1231

QY 662 TGAACGGGCGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGAGTCAACGAGA 721
Db 1232 TGAACGGGCGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGAGTCAACGAGA 1291
QY 722 ACCTGACCTCTCTGCTGAAGAGTCCACGCAAGAGTGAAGAGCTCTTTCAGGAGATCA 781
Db 1292 ATGTGA--CCTTGTGAAGAGTCCACGCAAGAGTGAAGAGCTCTTTCAGGAGATCA 1348
QY 782 CAGCTCTCTCTGCGCTCTCCATCTCATCAAACTGAACAGGAGACCGA--CCCTTGCCTG 840
Db 1349 CAGCTCTCTCTGCGCTCTCCATCTCATCAAACTGAACAGGAGACCGA--CCCTTGCCTG 1408
QY 841 TCGTGTCCAGGAATGTCAAGTGTCTGATGCCAAATGCAAAAAGGAGCGGAAGAAAAAGA 900
Db 1409 TCGTGTCCAGGAATGTCAAGTGTCTGATGCCAAATGCAAAAAGGAGCGGAAGAAAAAGC 1468
QY 901 AAGTGACCAACATAATCTCAATTTGATGATGAGGAAGATGAGCAGAACTCTGGGAGCTGT 960
Db 1469 AAGTGACCAACATAATCTCAATTTGATGATGAGGAAGATGAGCAGAACTCTGGGAGCTGT 1528
QY 961 TTAAGAGACACCTGGGGGAGGAGCTCAGAGGACAACTCCGACCGCTCTCTCTCTCA 1020
Db 1529 TTAAGAGACACCTGGGGGAGGAGCTCAGAGGACAACTCCGACCGCTCTCTCTCTCTCA 1588
QY 1021 ATATCATGTCCGCTTTTGAAGAGCCCTTCGGGCGCTTAACTCCAATGAGAA--TCAGAGCAGCA 1079
Db 1589 ATATCATGTCCGCTTTTGAAGAGCCCTTCGGGCGCTTAACTCCAATGAGAGTCAGAGCAGCA 1648
QY 1080 ACTCATGGAAATTTGATTCCTCTGCTTTTGAACGGGAGTTTGGGTACAGAGCTTGATG 1139
Db 1649 ACTCATGGAAATTTGATTCCTCTGCTTTTGAACAGGAGTTTGGGTACAGAGCTTGATG 1708
QY 1140 TGAAGACATCGATGATGAAGTGTGATGAAACGAGAGTACGCTGTATGGAACATCAT 1199
Db 1709 TGAAGACATCGATGATGAAGTGTGATGAAACGAGAGTACGCTGTATGGAACATCAT 1768
QY 1200 CAGGAAGGAGCAGCAGGGGCGCTCGGAGTCCGCCGAGAA 1239
Db 1769 CAGGAAGGAGCAGCAGGGGCGCTCGGAGTCCGCCGAGAA 1808

RESULT 14
US-10-917-503-18174
; Sequence 18174, Application US/10917503
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/10/917,503
; CURRENT FILING DATE: 2004-08-13
; PRIOR APPLICATION NUMBER: US/09/629,469
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
```

```

; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 13025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18174
; LENGTH: 6045
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (364)..(954)
US-10-917-503-18174

```

Query Match 41.6%; Score 792.8; DB 65; Length 6045;
Best Local Similarity 96.8%; Pred. No. 1.5e-184;
Matches 852; Conservative 0; Mismatches 22; Indels 6; Gaps 4;

363	AGAGCCCGTGTTCTTGTTACTACGTGAAGGAGGTCTCTCAACAAGCACAGACTGC-AGCGCT	421
932	AAAGCCCGTGTCTTGTTACTACGTGAAGGAGGTCTCTCAACAAGCACAGACTGC	991
422	TCTACTCCCTCGCGCCACATCGCCTCAGACGTGGGCGGGGTTCGGCCTGGCTGGCTGTG	481
992	TCTACTCCCTTCGCGCCACATCGCCTCAGACGTGGGCGGGGTTCGGCCTGGCTGTG	1051
482	CCCTCAACGAACACTCCCTCGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGC	541
1052	CCCTCAACGAACACTCCCTCGAGCGCTACCTGCACATGCTCCTGGCCGACCGCTGCAGGC	1111
542	TGAGCACTTTTATGAAGA CTGGTCTTTTGTGATGAGTGAAGAAAGTCCAGTATGCTTC	601
1112	TCAGTACTTTTATGAAGACTGGTCTTTTGTGATGAGTGAAGAGAGTCCAGTATGCTTC	1171
602	CTACCATGGCAGCAGGCTCGAACTCCATACTCTTTGCGATTATACATCGACAAACAAGATT	661
1172	CTACCATGGCAGCAGGCTCGAACTCCATACTCTTTGCGATTATACATCGACAAACAAGATT	1231
662	TGAACGGCGAGAGTAAGTTTGTCTCCCAACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA	721
1232	TGAACGGCGAGAGTAAGTTTGTCTCCCAACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA	1291
722	ACGTGACCTCCTTGTCTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCA	781
1292	ATGTGA--CCTTGTCAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTCAGGGAGATCA	1348
782	CAGCCTCCTCTGCGGCTCCTCATCTATACTCAAACTTGAAACAGGAGACCGA-CCTTTGCGTG	840
1349	CAGCCTCCTCTGCGCATCTCCATCTCATCAAACTTGAAACAGGAGACCGACCCCTTGCCCG	1408
841	TCTGTGCCGAATGTCAAGTGTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAAGA	900
1409	TCTGTGCCGAATGTCAAGTGTGATGCCAAATGCAAAAAGGAGCGGAAGAAGAAANG	1468
901	AAGTGACCAATAATCTCATTTTGATGATCAGGAAGATGAGCAGAACTCTGGGGACGTGT	960
1469	AAGTGACCAATAATCTCATTTTGATGATCAGGAAGATGAGCAGAACTCTGGGGACATGT	1528
961	TTAAAAAGACACCTGGGGCAGGGGAGCTCAGAGGACAACTCCGACCGTCTCTCTGTCTCA	1020
1529	TTAAAAAGACACCTGGGGCAGGGGAGCTCAGAGGACAACTCCGACCACTCTCTCTGTCTCA	1588
1021	ATATCATGTCCGCTTTGAAAGCCCTTTCGGGCTTACTCCAATGGAA-TCAGAGCAGCA	1079
1589	ATATCATGTCCGCTTTGAAAGCCCTTTCGGGCGAACTCCAATGGAAGTTCAGAGCAGCA	1648
1080	ACTCATGGAAAAATTGATTCCCTGCTTTGAAACGGGGAGTTTGGGTACCGAAGCTTGATG	1139
1649	ACTCGTGGAAAAATTGATTCCCTGCTTTGAAACAGGGAGTTTGGGTACCGAAGCTTGATG	1708
1140	TGAAAAAGCATCTGATGATGAAGATGTGGATGAAAAACGAAGATGACGTGTATGGAACCTCAT	1199

Mon Jun 13 13:27:09 2005

Qy	1035	TTTGAAGCCCTTCGGCCCTAACTCCAATGGAA-TCAGAGCAGCAACTCATGGAAAAATT	1093
Db	1222	TTTGAAGCCCTTCGGCCCAACTCCCATGGAAAGTCAGAGCAGCAACTCGTGGAAAAATT	1281
Qy	1094	GATTCCTGTCTTTGAACGGGAGTTTGGGTACCAGAAAGCTTGATGTGAAAAAGCATCGAT	1153
Db	1282	GATTCCTGTCTTTGAACAGGGAGTTTGGGTACCAGAAAGCTTGATGTGAAAAAGCATCGAT	1341
Qy	1154	GATGAAGATGTGGATGAAAAACGAAGATGACGTGTATGGAACTCATCAGGAAGGAAGCAC	1213
Db	1342	GATGAAGATGTGGATGAAAAACGAAGATGACGTGTATGGAACTCATCAGGAAGGAAGCAC	1401
Qy	1214	AGGGGCCACTCGAGTCGCCGAGAGCCACTGGAAGGGGAACACCTG	1260
Db	1402	AGGGGCCACTCAGAGTCGCCGAGAGTGCTGCCATGGCCACGACTG	1448

Search completed: June 11, 2005, 04:00:35
Job time : 4771.45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 09:44:05 ; Search time 729.11 Seconds
(without alignments)
11471.453 Million cell updates/sec

Title: US-09-155-676B-1

Perfect score: 1906

Sequence: 1 catgggtacgcgggtggcg.....tcattctcttgaaaaaaa 1906

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 12985247 seqs, 2194111492 residues

Total number of hits satisfying chosen parameters: 25970494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_NA_New.*

- 1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
- 9: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
- 10: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
- 11: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
- 12: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq.*
- 13: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq.*
- 14: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq.*
- 15: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1899	99.6	1906	6	US-09-155-676B-1
2	1143	60.0	2248	15	US-60-680-544-46181
3	1143	60.0	2248	15	US-60-680-473-46181
4	791	41.5	876	15	US-60-680-544-46660
5	791	41.5	876	15	US-60-680-473-46660
6	459	24.1	1487	1	PCT-US04-37982-81
7	376	19.7	527690	1	PCT-US04-37982-76
8	333.6	17.5	376	15	US-60-680-544-39116
9	333.6	17.5	376	15	US-60-680-473-39116
10	304.2	16.0	1725	1	PCT-US04-37982-72
11	293.4	15.4	1622	1	PCT-US04-37982-74
12	274.6	14.4	454	8	US-10-450-763-214
13	221	11.6	369941	1	PCT-US04-37982-71
14	206.8	10.8	465	13	US-11-060-756-3234
15	206.8	10.8	465	13	US-11-060-756-7506
16	195.6	10.3	463	8	US-10-450-763-28004
17	135.4	7.1	474	8	US-10-450-763-9353
18	119	6.2	911	8	US-10-450-763-215

19	106.2	5.6	331	12	US-11-124-617-290	Sequence 290, App
20	66.8	3.5	1796	12	US-11-097-143-11288	Sequence 11288, A
21	66.8	3.5	3796	12	US-11-097-143-11287	Sequence 11287, A
22	63.8	3.3	4118	12	US-11-090-997-1431	Sequence 1431, Ap
23	63.4	3.3	3177	15	US-60-680-544-45380	Sequence 45380, A
24	63.4	3.3	3177	15	US-60-680-473-45380	Sequence 45380, A
25	61.6	3.2	1253	6	US-09-155-676B-4	Sequence 4, Appli
26	61.6	3.2	1253	13	US-11-060-756-1553	Sequence 1553, Ap
27	61.6	3.2	1253	13	US-11-060-756-1554	Sequence 1554, Ap
28	61.6	3.2	1253	13	US-11-060-756-5825	Sequence 5825, Ap
29	61.6	3.2	1253	13	US-11-060-756-5826	Sequence 5826, Ap
30	59.4	3.1	5684	9	US-10-472-963-20	Sequence 20, Appl
31	54.8	2.9	36731	11	US-10-940-774A-13770	Sequence 13770, A
32	54.8	2.9	79977	2	PCT-US04-42189-58	Sequence 58, Appl
33	54.6	2.9	4153	12	US-11-090-997-343	Sequence 343, App
34	54.4	2.9	601	11	US-10-940-774A-69138	Sequence 69138, A
35	54.4	2.9	813	8	US-10-450-763-21166	Sequence 21166, A
36	54.4	2.9	188682	12	US-11-112-908-23	Sequence 23, Appl
37	52.4	2.7	537	15	US-60-680-544-41777	Sequence 41777, A
38	52.4	2.7	537	15	US-60-680-473-41777	Sequence 41777, A
39	52.2	2.7	421	12	US-11-000-688A-1575	Sequence 1575, Ap
40	52	2.7	1292	15	US-60-680-544-38736	Sequence 38736, A
41	52	2.7	1292	15	US-60-680-473-38736	Sequence 38736, A
42	52	2.7	1297	15	US-60-680-544-1851	Sequence 1851, Ap
43	52	2.7	1297	15	US-60-680-473-1851	Sequence 1851, Ap
44	50.6	2.7	16420	11	US-10-990-328A-94349	Sequence 94349, A
45	50	2.6	9811	11	US-10-940-774A-12759	Sequence 12759, A

ALIGNMENTS

RESULT 1

US-09-155-676B-1
Sequence 1 Application US/09155676B
GENERAL INFORMATION:
APPLICANT: WALLACH, David
APPLICANT: MALININ, Nikolai
APPLICANT: BOLDIN, Mark
APPLICANT: KOVALENKO, Andrei
APPLICANT: METT, Igor
TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR
FILE OF INVENTION: PREPARATION AND USE
FILE REFERENCE: WALLACH=21
CURRENT APPLICATION NUMBER: US/09/155,676B
CURRENT FILING DATE: 1999-01-04
PRIOR APPLICATION NUMBER: PCT/IL97/00117
PRIOR FILING DATE: 1997-04-01
PRIOR APPLICATION NUMBER: IL 117800
PRIOR FILING DATE: 1996-04-02
PRIOR APPLICATION NUMBER: IL 119133
PRIOR FILING DATE: 1996-08-26
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin version 3.3
SEQ ID NO 1
LENGTH: 1906
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (94)..(94)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (110)..(110)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (115)..(115)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (129)..(129)

OTHER INFORMATION: n is a, c, g, or t									
FEATURE: NAME/KEY: misc_feature									
LOCATION: (131)..(131)									
OTHER INFORMATION: n is a, c, g, or t									
FEATURE: NAME/KEY: misc_feature									
LOCATION: (136)..(136)									
OTHER INFORMATION: n is a, c, g, or t									
FEATURE: NAME/KEY: misc_feature									
LOCATION: (202)..(202)									
OTHER INFORMATION: n is a, c, g, or t									
US-09-155-676B-1									
Query Match									
Best Local Similarity 99.6%; Score 1899; DB 6; Length 1906;									
Matches 1906; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	CATTGGGTACCGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCAGGAATTCG	60						
Db	1	CATTGGGTACCGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCAGGAATTCG	60						
Qy	61	ATTGAGGCGCAGAAAGCGCGCGCGCGCGCANGCACCGCGCGGGANAGGNCNCAATG	120						
Db	61	ATTGAGGCGCAGAAAGCGCGCGCGCGCGCANGCACCGCGCGGGANAGGNCNCAATG	120						
Qy	121	ACGGGATCNCNCAATGACAAAAGACAAATTTCTGTGGAGCGACTCTCGATGTCAGTG	180						
Db	121	ACGGGATCNCNCAATGACAAAAGACAAATTTCTGTGGAGCGACTCTCGATGTCAGTG	180						
Qy	181	AAACAGTGCAGATCCGCTTTNGAGGAGAAAGAGATGCTCGATTCCGACAGCAGG	240						
Db	181	AAACAGTGCAGATCCGCTTTNGAGGAGAAAGAGATGCTCGATTCCGACAGCAGG	240						
Qy	241	GTACCTGTCTGTGCCAGTTTGAAGCCGTCCTGACGATGGCTTGAAGAGGAGTCTGA	300						
Db	241	GTACCTGTCTGTGCCAGTTTGAAGCCGTCCTGACGATGGCTTGAAGAGGAGTCTGA	300						
Qy	301	GGATTGGCACTACAGCGCGCAGATCAAGCAGGAGCGGGCTTGGCCAGAAACCCGAA	360						
Db	301	GGATTGGCACTACAGCGCGCAGATCAAGCAGGAGCGGGCTTGGCCAGAAACCCGAA	360						
Qy	361	ACAGAGCCGTGTTCTGGTACTACGTGAAGAGGTCTCTCAACAGCAGCAGCTGCAGCGC	420						
Db	361	ACAGAGCCGTGTTCTGGTACTACGTGAAGAGGTCTCTCAACAGCAGCAGCTGCAGCGC	420						
Qy	421	TTCTACTCCCTGCGCCACATCGCCTCAGACGTGGCGCGGGTTCGGCTGGCTGT	480						
Db	421	TTCTACTCCCTGCGCCACATCGCCTCAGACGTGGCGCGGGTTCGGCTGGCTGT	480						
Qy	481	GCCCTCAACGAAACATCTCTGGAGCGCTACCTGCATGCTCTGGCCGACCGCTGCAGG	540						
Db	481	GCCCTCAACGAAACATCTCTGGAGCGCTACCTGCATGCTCTGGCCGACCGCTGCAGG	540						
Qy	541	CTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTT	600						
Db	541	CTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTT	600						
Qy	601	CCTTACCATGGCAGCAGTCTGAATCCATCTCTTTGGCATTAACATCGACAAACAGGAT	660						
Db	601	CCTTACCATGGCAGCAGTCTGAATCCATCTCTTTGGCATTAACATCGACAAACAGGAT	660						
Qy	661	TTGAACGGGAGAGTAAGTTTGTCTCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCGAG	720						
Db	661	TTGAACGGGAGAGTAAGTTTGTCTCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCGAG	720						
Qy	721	AACGTGACCTCTTGTCTGAAGAGTCCACGCAAGAGTGAAGACCTGTTTCAGGAGATC	780						
Db	721	AACGTGACCTCTTGTCTGAAGAGTCCACGCAAGAGTGAAGACCTGTTTCAGGAGATC	780						
Qy	781	ACAGCTCTCTGCGGTCTCCTCTCATCAAACTGAACAGGAGACCGACCTTGCTG	840						
Db	781	ACAGCTCTCTGCGGTCTCCTCTCATCAAACTGAACAGGAGACCGACCTTGCTG	840						

Db	781	ACAGCTCTCTGCGGTCTCCTCATCAAACTGAACAGGAGACCGACCTTGCTG	840
Qy	841	TCGTGTCCAGGAATGTCTAGTGTGCTGATGCCAATGCAAAAAGCAGCGGAAGAAAAGA	900
Db	841	TCGTGTCCAGGAATGTCTAGTGTGCTGATGCCAATGCAAAAAGCAGCGGAAGAAAAGA	900
Qy	901	AAGTGACCAACATAATCTCATTTGATGATGAGGAGATGAGCAACCTCTCGGGACGTGT	960
Db	901	AAGTGACCAACATAATCTCATTTGATGATGAGGAGATGAGCAACCTCTCGGGACGTGT	960
Qy	961	TTAAAAAGACACCTCGGGAGGAGCTCAGAGCAACCTCCGACCGCTCTCTGTCA	1020
Db	961	TTAAAAAGACACCTCGGGAGGAGCTCAGAGCAACCTCCGACCGCTCTCTGTCA	1020
Qy	1021	ATATCATGTCCGCTTTTGAAGCCCTTCGGGCCCTAACTCCAATGGAATCAGAGCAGAA	1080
Db	1021	ATATCATGTCCGCTTTTGAAGCCCTTCGGGCCCTAACTCCAATGGAATCAGAGCAGAA	1080
Qy	1081	CTCATCGAAAAATTGATTCCTCTCTTTGAAACCGGGAGTTTGGGTACCGAAGCTTGATGT	1140
Db	1081	CTCATCGAAAAATTGATTCCTCTCTTTGAAACCGGGAGTTTGGGTACCGAAGCTTGATGT	1140
Qy	1141	GAAAAATCGATGATGAAGATGTGGATGAAACCGAAGATGACGTGTATGGAATCTCATC	1200
Db	1141	GAAAAATCGATGATGAAGATGTGGATGAAACCGAAGATGACGTGTATGGAATCTCATC	1200
Qy	1201	AGGAAGAAAGCACAGCGGCCACTCGGAGTCGCGCGAGAAAGCCACTGGAAGGAAACCTGT	1260
Db	1201	AGGAAGAAAGCACAGCGGCCACTCGGAGTCGCGCGAGAAAGCCACTGGAAGGAAACCTGT	1260
Qy	1261	CCTCTCCAGATGACAGCTGGGCTCCGCTGAAGTGTCTGCACATGACTCCGACATCT	1320
Db	1261	CCTCTCCAGATGACAGCTGGGCTCCGCTGAAGTGTCTGCACATGACTCCGACATCT	1320
Qy	1321	CTTCCCTGTCAAGTGGGTCTCTACAGCCAGCAGATGCCCCCTCGGAAGCCTGGA	1380
Db	1321	CTTCCCTGTCAAGTGGGTCTCTACAGCCAGCAGATGCCCCCTCGGAAGCCTGGA	1380
Qy	1381	GAAAGGACAGGACCAAGAGACCAAGTTCCTCCGATCTCTGGACTTCGGTACAGTGTGA	1440
Db	1381	GAAAGGACAGGACCAAGAGACCAAGTTCCTCCGATCTCTGGACTTCGGTACAGTGTGA	1440
Qy	1441	AGCCAGCTCTCCAGGCGACCGAAGTCTCTGACAGCTCTTACTTCTGCTCAGTGCCA	1500
Db	1441	AGCCAGCTCTCCAGGCGACCGAAGTCTCTGACAGCTCTTACTTCTGCTCAGTGCCA	1500
Qy	1501	GAGTCCATGACAAATTAGTGAATCTGCGCCAGCCACTGTGCGCATGATGAACAGGAAGAT	1560
Db	1501	GAGTCCATGACAAATTAGTGAATCTGCGCCAGCCACTGTGCGCATGATGAACAGGAAGAT	1560
Qy	1561	GAGTGGAGGAGGAGAACAGATCACTCGAATCTGCGAATCTGCTGACCGTGAAGTGGACCTCA	1620
Db	1561	GAGTGGAGGAGGAGAACAGATCACTCGAATCTGCGAATCTGCTGACCGTGAAGTGGACCTCA	1620
Qy	1621	GCCGCGTCCGCGCAAGAGTGGACACCTTGAAGAGGAGTGGCTGAACAGGAGGAGCGG	1680
Db	1621	GCCGCGTCCGCGCAAGAGTGGACACCTTGAAGAGGAGTGGCTGAACAGGAGGAGCGG	1680
Qy	1681	CAGGCGATGAAGGTCCAGGCGCTGCGCCAGCTATCTTTGCTATTTTGTGAGGAGATCTAA	1740
Db	1681	CAGGCGATGAAGGTCCAGGCGCTGCGCCAGCTATCTTTGCTATTTTGTGAGGAGATCTAA	1740
Qy	1741	CCCCACGTGAGAACCATGCTGGTGGAGAAATCGAGGAGAGAGAAATCAACAGTTCCTGA	1800
Db	1741	CCCCACGTGAGAACCATGCTGGTGGAGAAATCGAGGAGAGAGAAATCAACAGTTCCTGA	1800
Qy	1801	TAGTCTCATTTTGTGCTCTCTGATCCAGTCTTTCCTGAAGCTGTGTTCCTCTGGACCTTT	1860
Db	1801	TAGTCTCATTTTGTGCTCTCTGATCCAGTCTTTCCTGAAGCTGTGTTCCTCTGGACCTTT	1860
Qy	1861	CATGTATGTGAGCCCAATAAATTGCTTTTCATTCCTTGAAGAAAAAAA	1906
Db	1861	CATGTATGTGAGCCCAATAAATTGCTTTTCATTCCTTGAAGAAAAAAA	1906

Db	622	AGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGCAAGACGTGACCTCCTTGC	681
Qy	737	TGAAGGAGTCCACGCAAGGAGTGAAGAGTCTGTTTCAGGGAGATCACAGCCTCCTCTGCCG	796
Db	682	TGAAGGAGTCCACGCAAGGAGTGAAGAGTCTGTTTCAGGGAGATCACAGCCTCCTCTGCCG	741
Qy	797	TCTCCATCTCATCAAACTTGAACAGGAGACCGA - CCCTTGCCTGTCTGTCCAGGAATG	855
Db	742	TCTCCATCTCATCAAACTTGAACAGGAGACCGA - CCCTTGCCTGTCTGTCCAGGAATG	801
Qy	856	TCAGTGTCTGATGCCAAATGCAAAAAGGAGCGGAAGAAAGAAAGTGCACCAACATAA	915
Db	802	TCAGTGTCTGATGCCAAATGCAAAAAGGAGCGGAAGAAAGAAAGTGCACCAACATAA	861
Qy	916	TCTCATTTGATGATGAGGAAGATGAGCAGAACTCTCGGGACGTGTTTAAAAAGACACCTG	975
Db	862	TCTCATTTGATGATGAGGAAGATGAGCAGAACTCTCGGGACGTGTTTAAAAAGACACCTG	921
Qy	976	GGGACGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTGTCTCAATATCATGTCCGCCT	1035
Db	922	GGGACGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTGTCTCAATATCATGTCCGCCT	981
Qy	1036	TTGAAAGCCCCCTTCGGSCCTTAATCCAAATCGAA - TCAGACGAGCAACTCATGGAAATATG	1094
Db	982	TTGAAAGCCCCCTTCGGSCCTTAATCCAAATCGAAAGTCAAGCAGCACTCATGGAAATATG	1041
Qy	1095	ATTCCCTGTCTTTGAACGGGGAGTTTGGGTACAGAAAGCTTGATGTGAAAGCATCGATG	1154
Db	1042	ATTCCCTGTCTTTGAACGGGGAGTTTGGGTACAGAAAGCTTGATGTGAAAGCATCGATG	1101
Qy	1155	ATGAAGATGTGGATGAAACGAAAGATGACGTGTATGGAAACTCATCAGGAAGGAGCACA	1214
Db	1102	ATGAAGATGTGGATGAAACGAAAGATGACGTGTATGGAAACTCATCAGGAAGGAGCACA	1161
Qy	1215	GGGGCCACTCCGAGTCCCGCAGGAAGCCACTGGAAGGGAACACCTCGCTCTCCAGATGC	1274
Db	1162	GGGGCCACTCCGAGTCCCGCAGGAAGTAAAGTTGTGTGTGAAGTGGAGTCTCACCGTCCC	1221
Qy	1275	ACAGCTGGGCTCCGCTGAAGGTGCTGCACAAATGACTCCGACATCCTCTTCCCTG	1328
Db	1222	CCAGGCTGGAGTGCCTGGCGGTGATCTCAGCTCACTGCAACCTCCACCTCCCGG	1275

Db 475 GTGTCAGGAATGTGCTGCTGATGCCAAATGCAAAAGGAGCGGAAGAAAGCA 534
Qy 903 GTGACCAATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAGCTGTTT 962
Db 535 GTGACCAACATTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAGCTGTTT 594
Qy 963 AAAAAAGACACTGGGGCAGGGGAGAGCTCAGAGACAACTCCGACCGCTCTCTGTCAAT 1022
Db 595 AAAAAAGACACTGGGGCAGGGGAGAGCTCAGAGACAACTCCGACCGCTCTCTGTCAAT 654
Qy 1023 ATCATGTCCGCTTTGAAAGCCCTTCGGGCTAACTCCAAATGAA--TCAGAGCAGCAAC 1081
Db 655 ATCATGTCCGCTTTGAAAGCCCTTCGGGCTAACTCCAAATGAAAGTCAAGAGCAGCAAC 714
Qy 1082 TCATGGAATAATGATTCCTCTGCTTTGAAACGGGAGTTCGGGTACCAAGAGCTTGATGTG 1141
Db 715 TCGTGGAAATGATTCCTCTGCTTTGAAACAGGAGTTCGGGTACCAAGAGCTTGATGTG 774
Qy 1142 AAAAGCATCGATGATGAAGATGTGGATGAAACGAAAGATGACGTGTATGAAACTCATCA 1201
Db 775 AAAAGCATCGATGATGAAGATGTGGATGAAACGAAAGATGACGTGTATGAAACTCATCA 834
Qy 1202 GGAAGGAGCAGCAGGGGCCACTCGAGTCGCCCGAGAG 1240
Db 835 GGACGGAAGCAGCAGGGGCCACTCGAGTCGCCCGAGAG 873

RESULT 5

US-60-680-473-46660
; Sequence 46660, Application US/60680473
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary to
; TITLE OF INVENTION: Fragments of, Cynomolgus Monkey Genes and the Use Thereof
; FILE REFERENCE: 21590290000
; CURRENT APPLICATION NUMBER: US/60/680,473
; CURRENT FILING DATE: 2005-05-13
; NUMBER OF SEQ ID NOS: 48714
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 46660
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-60-680-473-46660

Query Match 41.5%; Score 791; DB 15; Length 876;
Best Local Similarity 96.8%; Pred. No. 1.3e-176;
Matches 851; Conservative 0; Mismatches 20; Indels 8; Gaps 4;
Qy 364 GAGCCGCTGTCTGGTACTACGTGAAGAGGTCTCAACAGCAGCAGCTGACGCGTTC 423
Db 1 GAGCCGCTGTCTGGTACTACGTGAAGAGGTCTCAACAGCAGCAGCTGACGCGTTC 57
Qy 424 TACTCCCTGCGCCACATCGCCTCAGACGTGGCGGGGTTCGCGCTGCGCTGTGTC 483
Db 58 TACTCCCTGCGCCACATCGCCTCAGACGTGGCGGGGTTCGCGCTGCGCTGTGTC 117
Qy 484 CTCAAGCAACACTCCCTGGAGCGCTACCTGCAATGCTCTGCGCCAGCAGCTGAGCGTG 543
Db 118 CTCAAGCAACACTCCCTGGAGCGCTACCTGCAATGCTCTGCGCCAGCAGCTGAGCGTG 177
Qy 544 AGCACTTTTATGAGACTGGTCTTTGTGATGATGAGAAAGTCCAGATGCTTCCT 603
Db 178 AGTACTTTTATGAGACTGGTCTTTGTGATGATGAGAAAGTCCAGATGATGCTTCCT 237

Qy 604 ACCATGTCAGCAGGTCTGAACCTCATATCTTTTCCGATTAAACATCGAACAGGATTTG 663
Db 238 ACCATGTCAGCAGGTCTGAACCTCATATCTTTTCCGATTAAACATCGAACAGGATTTG 297
Qy 664 AACGGCAGAGTAAGTTTGTCTCCACCGTTTCAAGCTCTTAAAGGAGTCAACGCAAG 723
Db 298 AACGGCAGAGTAAGTTTGTCTCCACCGTTTCAAGCTCTTAAAGGAGTCAACGCAAG 357
Qy 724 GTGACCTCTCTGTGAAGAGTCCAGCAAGGAGTGAAGAGCTGTTTTCAGGGAGTCA 783
Db 358 GTGA---CCTTGTGAAGAGTCCAGCAAGGAGTGAAGAGCTGTTTTCAGGGAGTCA 414
Qy 784 GCCTCTCTGCTGCTCTCCATCTCATCAAACTGAAACAGAGACCGA-CCTTGGCTGTG 842
Db 415 GCCTCTCTGCTGCTCTCCATCTCATCAAACTGAAACAGAGACCGA-CCTTGGCTGTG 474
Qy 843 GTCTCAGGAAATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 902
Db 475 GTCTCAGGAAATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 534
Qy 903 GTGACCAACATAATCTCATTTGATGATGAGAAAGATGAGCAGAACTCTGGGGAGCTGTTT 962
Db 535 GTGACCAACATAATCTCATTTGATGATGAGAAAGATGAGCAGAACTCTGGGGAGCTGTTT 594
Qy 963 AAAAAAGACACTGGGGCAGGGGAGAGCTCAGAGCAAACTCCGACCGCTCTCTGTCAAT 1022
Db 595 AAAAAAGACACTGGGGCAGGGGAGAGCTCAGAGCAAACTCCGACCGCTCTCTGTCAAT 654
Qy 1023 ATCATGTCCGCTTTGAAAGCCCTTCGGGCTAACTCCAAATGAA--TCAGAGCAGCAAC 1081
Db 655 ATCATGTCCGCTTTGAAAGCCCTTCGGGCTAACTCCAAATGAAAGTCAAGAGCAGCAAC 714
Qy 1082 TCATGGAATAATGATTCCTCTGCTTTGAAACGGGAGTTCGGGTACCAAGAGCTTGATGTG 1141
Db 715 TCGTGGAAATGATTCCTCTGCTTTGAAACAGGAGTTCGGGTACCAAGAGCTTGATGTG 774
Qy 1142 AAAAGCATCGATGATGAAGATGTGGATGAAACGAAAGATGACGTGTATGAAACTCATCA 1201
Db 775 AAAAGCATCGATGATGAAGATGTGGATGAAACGAAAGATGACGTGTATGAAACTCATCA 834
Qy 1202 GGAAGGAGCAGCAGGGGCCACTCGAGTCGCCCGAGAG 1240
Db 835 GGACGGAAGCAGCAGGGGCCACTCGAGTCGCCCGAGAG 873

RESULT 6

PCT-US04-37982-81
; Sequence 81, Application PC/TUS0437982
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS IN CANCER
; FILE REFERENCE: CHIR0062-500 (23356.0003)
; CURRENT APPLICATION NUMBER: PCT/US04/37982
; CURRENT FILING DATE: 2004-10-21
; PRIOR APPLICATION NUMBER: US 10/692,382
; PRIOR FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: US 10/691,209
; PRIOR FILING DATE: 2003-10-21
; NUMBER OF SEQ ID NOS: 3504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US04-37982-81

Query Match 24.1%; Score 459; DB 1; Length 1487;
Best Local Similarity 99.8%; Pred. No. 3.4e-98;
Matches 470; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 1240 GCCACTGGAAGGGAACACCTGCTCTCCAGATGCAAGCTGGCTCCGCTGAAGTGCT 1299

Db	1	GC	CA	CT	TG	GA	AG	GA	AA	CA	CT	GC	CT	CT	CC	CA	GA	TG	CA	CA	GC	TG	GG	CT	CC	GC	T	GA	AG	GT	GCT	60
Qy	1300	GC	CA	AT	GA	CT	CG	GA	CA	T	CT	T	T	CC	CT	GT	CAG	TG	CG	TG	GG	CT	CT	T	CA	CAG	CC	CA	G	AGA	1359	
Db	61	GC	CA	AT	GA	CT	CG	GA	CA	T	CT	T	CC	CT	GT	CAG	TG	CG	TG	GG	CT	CT	T	CA	CAG	CC	CA	G	AGA	120		
Qy	1360	TG	CC	CC	CC	CT	CG	GA	AG	CC	TT	GG	AG	AA	CG	CG	GA	CA	GA	GA	CA	CA	CT	T	TC	CC	GG	GA	T	CC	1419	
Db	121	TG	CC	CC	CC	CT	CG	GA	AG	CC	TT	GG	AG	AA	CG	CG	GA	CA	GA	GA	CA	CA	CT	T	TC	CC	GG	GA	T	CC	180	
Qy	1420	TG	GA	CT	TC	GG	TA	CAG	TG	TG	GA	AG	CC	AG	CT	CT	CA	GG	CC	CA	CG	GA	GT	CC	CT	CA	AG	CA	GC	CT	1479	
Db	181	TG	GA	CT	TC	GG	TA	CAG	TG	TG	GA	AG	CC	AG	CT	CT	CA	GG	CC	CA	CG	GA	GT	CC	CT	CA	AG	CA	GC	CT	240	
Qy	1480	GT	TA	-	CT	T	GC	CT	CAG	TG	CC	AG	T	CC	AT	GA	CT	GT	GA	CT	GT	GA	CT	GT	GA	CT	GT	GA	CT	GT	1538	
Db	241	GT	TA	CT	T	GC	CT	CAG	TG	CC	AG	T	CC	AT	GA	CT	GT	GA	CT	GT	GA	CT	GT	GA	CT	GT	GA	CT	GT	300		
Qy	1539	GG	CA	CT	GA	TA	GA	CA	AG	GA	AG	TA	GA	CT	GG	AG	GA	GA	CA	GA	TA	CA	CT	TG	GA	AA	CA	CT	GC	T	1598	
Db	301	GG	CA	CT	GA	TA	GA	CA	AG	GA	AG	TA	GA	CT	GG	AG	GA	GA	CA	GA	TA	CA	CT	TG	GA	AA	CA	CT	GC	T	360	
Qy	1599	CG	AG	GT	GA	AT	TG	AG	CA	CT	CAG	CC	GG	CT	CC	GA	AG	GT	GA	CA	CA	CT	TG	GA	AA	AG	GA	1658				
Db	361	CG	AG	GT	GA	AT	TG	AG	CA	CT	CAG	CC	GG	CT	CC	GA	AG	GT	GA	CA	CA	CT	TG	GA	AA	AG	GA	420				
Qy	1659	GG	T	GG	CT	GA	AC	CA	GG	AG	GA	CG	CA	GG	CA	TG	AA	GG	TA	GG	TA	CT	CA	GG	CG	T	GG	CC	AG	1709		
Db	421	GG	T	GG	CT	GA	AC	CA	GG	AG	GA	CG	CA	GG	CA	TG	AA	GG	TA	GG	TA	CT	CA	GG	CG	T	GG	CC	AG	471		

```

RESULT 7
PCT-US04-37982-76
; Sequence 76, Application PC/TUS0437982
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS IN CANCER
; FILE REFERENCE: CHIR0062-500 (23356.0003)
; CURRENT APPLICATION NUMBER: PCT/US04/37982
; CURRENT FILING DATE: 2004-10-21
; PRIOR APPLICATION NUMBER: US 10/692,382
; PRIOR FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: US 10/691,209
; PRIOR FILING DATE: 2003-10-21
; NUMBER OF SEQ ID NOS: 3504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 527690
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(527690)
; OTHER INFORMATION: n = A, C, G or T/U
PCT-US04-37982-76

```

	Query Match	19.7%	Score 376;	DB 1;	Length 527690;
	Best Local Similarity	89.1%	Pred. No. 4.8e-78;		
	Matches 417;	Conservative 0;	Mismatches 50;	Indels 1;	Gaps 1;
Qy	862	CTGATGCCAAATGCAAAAAGGAGCGGAGAGAAAAGAAAGTAGTGACCAACATATATCTCAT	921		
Db	317	CAGATGCCAAATGCAAAAAGGAGCGGAGAGAAAAGAAAGTAGTGACCAACATATATCTCAT	376		
Qy	922	TTGATGATGAGGAAGATGAGCAGAACTCTTGGGGACGTGTTTAAAAAAGACACCTTGGGGCAG	981		
Db	377	TTGATGATGAGGAAGATGAGCAGAACTCTTGGGGACGTGTTTAAAAAAGACACCTTGGGGCAG	436		
Qy	982	GGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTGTCATATATCATGTCCGCCCTTTGAAA	1041		
Db	437	GGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTGTCATATATCATGTCCGCCCTTTGAAA	496		

Qy	1042	GCCTTCGGGCTAACTCCAATGGAA-TCAGAGCAGCAACTCATCGAAAAATTGATTCCC	1100
Db	497	GCCTTCGGGCTAACTCCAATGGAAATCAGAGCAGCAACTCATCGAAAAATTGATTCCC	556
Qy	1101	TGTCCTTTGAACGGGGAGTTTGGGTACCAAGACCTTCGATGTGAAAAACATCGATGATGAAG	1160
Db	557	TGTCCTTTGAACGGGGAGTTTGGGTACCAAGACCTTCGATGTGAAAAACATCGATGATGAAG	616
Qy	1161	ATGTGGATGAAAACGAAGATGACGTGTATGGAAATCTCATCAGAGGAAGACACAGGGGCC	1220
Db	617	ATGTGGATGAAAACGAAGATGACGTGTATGGAAATCTCATCAGAGGAAGACACAGGGGCC	676
Qy	1221	ACTCGAGTCGCCCGAGAGCCACTGGAAGGAACACCTCGCTCTCCAGATGCACAGCT	1280
Db	677	ACTCGAGTCGCCCGAGAGTAAGTTTGTGTGAAGGTGGAGTCTCACCCTCCCCCAGGC	736
Qy	1281	GGGCTCCGCTGGAAGGTGCTGCACAAATGACTCCGACATCTCTTCCCCTG	1328
Db	737	TGAGATGCTCGTGGCTGATCTCAGCTCAGTCAACCTCCACCTCCCGG	784

RESULT 8

```

RES001. 8
US-60-680-544-39116
; Sequence 39116, Application US/60680544
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary to
; TITLE OF INVENTION: Fragments of, Cynomolgus Monkey Genes and the Use Thereof
; FILE REFERENCE: 21590290000
; CURRENT APPLICATION NUMBER: US/60/680,544
; CURRENT FILING DATE: 2005-05-13
; NUMBER OF SEQ ID NOS: 48714
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 39116
; LENGTH: 376
; TYPE: DNA
; ORGANISM: Macaca Mulatta
US-60-680-544-39116

```

Query Match 17.5%; Score 333.6; DB 15; Length 376;
Best Local Similarity 94.7%; Pred. No. 1e-68;
Matches 356; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

Qy	864	GATGCCAAATGCAAAAGAGGACGGAGAGAGAAAAGAAAGTAGACCAACATAATCTCATTT	923
Db	1	GATGCCAAATGCAAAAGAGGACGGAGAGAGAAAAGAAAGTAGACCAACATAATCTCATTT	60
Qy	924	GATGATGAGAAAGATGAGGAGAACTCTGGGGACGTCTTTAAAAGACACACTGGGGCAGGG	983
Db	61	GATGATGAGGAGATGAGCAAACTCTCGGGACATTTTAAAAGACACACTGGGGCAGGG	120
Qy	984	GAGAGCTCAGAGACAACTCCGACCGCTCCTCTGTCTCAATATCATGTCCGCTCTTGAAAGC	1043
Db	121	GAGAGCTCGGAGACAACTCCGACCGCTCCTCTGTAAATATATATCAGCCTTTGAAAGC	180
Qy	1044	CCCTTCGGGCTTAATCCAAATGGAA-TCAGAGCAGCAACTCATGGAAAAATTGATTCCTGT	1102
Db	181	CCCTTTGGGCCAAATTCCAATGGAAGTCAGAGCAGCAACTCGTGGAAAAATTGATTCCTGT	240
Qy	1103	TCCTTTGAACGGGGAGTTTGGGTACCGAAGCTTGATGTGAAAACGATCGATGTAAGAT	1164
Db	241	TCCTTTGAATGGGGAGTTTGGGTACCGAAGCTTGATGTGAAAACGATCGATGTAAGAT	300
Qy	1163	GTGATGAAAACGAAAGATGACGTGTATGAAAACTCATCAGGAAGGAAGCACAGGGGCCAC	1222
Db	301	GTGATGAAAACGAAAGATGACGTGTATGAAAACTCATCAGGAAGGAAGCACAGGGGCCAC	360

Qy 1223 TCGAGTCCGCCGAGA 1238
Db 361 GCAGAGTCACTGAGA 376

RESULT 9

US-60-680-473-39116
; Sequence 39116, Application US/60680473
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary to
; FILE OF INVENTION: Fragments of, Cynomolgus Monkey Genes and the Use Thereof
; FILE REFERENCE: 21590290000
; CURRENT APPLICATION NUMBER: US/60/680,473
; NUMBER OF SEQ ID NOS: 48714
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 39116
; LENGTH: 376
; TYPE: DNA
; ORGANISM: Macaca Mulatta
US-60-680-473-39116

Query Match 17.5%; Score 333.6; DB 15; Length 376;
Best Local Similarity 94.7%; Pred. No. 1e-68;
Matches 356; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

Qy 864 GATGCCAAATGCCAAAGGAGCGGAGAGCAAGAAAAAGAAAGTACCAACATATCTCATTT 923
Db 1 GATGCCAAATGCCAAAGGAGCGGAGAGCAAGAAAAAGAAAGTACCAACATATCTCATTT 60
Qy 924 GATGATGAGGAAGATGAGCAGAACTCTGGGACGCTGTTTAAAAAGACACCTGGGGCAGGG 983
Db 61 GATGATGAGGAGATGAGCAGAACTCTGGGACATTTTAAAAAGACACCTGGGGCAGGG 120
Qy 984 GAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAATATCATGTCGCGCTTTGAAAGC 1043
Db 121 GAGAGCTCGAGGACAACTCCGACCGCTCTCTGTAAATATAATATCAGCCTTTGAAAGC 180
Qy 1044 CCCTTGGGCTTAATCCAAATGGAA-TCAGAGCAGCAACTCATGGAATAATGATTCCTG 1102
Db 181 CCCTTGGGCTTAATCCAAATGGAAATGATGGAAGTACAGCAGCAACTCGTGGAAATGATTCCTG 240
Qy 1103 TCTTTGAACGGGAGTTTGGGTACCAAGCTTGATGTGAAAGCATCGATGATGAAGAT 1162
Db 241 TCTTTGAATGGGAGTTTGGGTACCAAGCTTGATGTGAAAGCATCGATGATGAAGAT 300
Qy 1163 GTGGATGAAACGAAGATGAGTGTATGGAACCTCATCAGGAAGGAGCAAGCGGCCAC 1222
Db 301 GTGGATGAAACGAAGATGAGTGTATGGAACCTCATCAGGAAGGAGCAAGCGGCCAC 360
Qy 1223 TCGAGTCCGCCGAGA 1238
Db 361 GCAGAGTCACTGAGA 376

RESULT 10

PCT-US04-37982-74
; Sequence 74, Application PC/TUS0437982
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS IN CANCER
; FILE REFERENCE: CHIR0062-500 (23356.0003)
; CURRENT APPLICATION NUMBER: PCT/US04/37982

; CURRENT FILING DATE: 2004-10-21
; PRIOR APPLICATION NUMBER: US 10/692,382
; PRIOR FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: US 10/691,209
; PRIOR FILING DATE: 2003-10-21
; NUMBER OF SEQ ID NOS: 3504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Mus musculus
PCT-US04-37982-74

Query Match 16.0%; Score 304.2; DB 1; Length 1725;
Best Local Similarity 80.1%; Pred. No. 1.2e-61;
Matches 387; Conservative 0; Mismatches 83; Indels 13; Gaps 2;

Qy 1240 GCCACTGGAGGGAACACCTGCTCCCGAGATGCACAGCTGGCTGGCTCCGCTGAAGTGCT 1299
Db 241 GACACTGGAGGGAACACCTGCTCCCGAGATGCACAGCTGGCTGGCTCCGCTGAAGTGCT 300
Qy 1300 GCACAA-----TGACTCCGACATCTCTTCCCTGTGAGTGGCTGGCTCCTA 1347
Db 301 TCATGAGATGCGGATGCTGACACTGATGTACTTCTCCCTGTGAGTGGCTGGCTCCTA 360
Qy 1348 CAGCCCGAGATGCCCCCTCGAAGCCTTGGAGAACCGGACAGGACCCAGAGGACACGCT 1407
Db 361 TGGTGACGACAGATGCCCTGTGGAGAGTCTGGAGAGTGGAGAACAGGAAACCATAT 420
Qy 1408 TCTCCCGGATCTGGACTTCGGTACAGTGTGGAGAGCAGCTCTCCAGGCCACGGAAGTCC 1467
Db 421 AATCCCGAGAGCTGGCTTCGGTACAGTGGAGAGCAGCTCTCCAGGCCACGGAAGTCC 480
Qy 1468 TCTGACGAGCTGTGA-CTTCTGCTCAGTGCAGAGTCCATGACAAATTAGTGAACCTGCG 1526
Db 481 TCTGACGAGCTGTGACTACTTCTGCTCAGTGCCTGAGTCCATGACAGTCCATGACTGCG 540
Qy 1527 CCAGGCCACTGTGGCCATGATGAACAGGAGGATGAGCTGGAGGAGGAGAACAGATCACT 1586
Db 541 CCAAGCCATCGTGGCCATGATGAACAGGAGGATGAATTGGAAGAGAGAGTAATGGTCACT 600
Qy 1587 GCGAAACCTGCTCGAGGAGATGAGGACACTCAGCGGCTCCGCGCAGAGGTGGACAC 1646
Db 601 GCGAAATCTCTTACCGGAGATGAGGACACTCAGCAGCGCTCCGCGCAGAGGTGGATGC 660
Qy 1647 CTTGAAAAAGAGGTGGCTGAACAGGAGGAGCGGAGGAGCATCAAGTCCAGGCGCTGGC 1706
Db 661 CTTGAGGCGGAGGTGACAGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
Qy 1707 CAG 1709
Db 721 CAG 723

RESULT 11

PCT-US04-37982-72
; Sequence 72, Application PC/TUS0437982
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS IN CANCER
; FILE REFERENCE: CHIR0062-500 (23356.0003)
; CURRENT APPLICATION NUMBER: PCT/US04/37982
; CURRENT FILING DATE: 2004-10-21
; PRIOR APPLICATION NUMBER: US 10/692,382
; PRIOR FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: US 10/691,209
; PRIOR FILING DATE: 2003-10-21
; NUMBER OF SEQ ID NOS: 3504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 1622
; TYPE: DNA

```

; ORGANISM: Mus musculus
PCT-US04-37982-72

Query Match      15.4%; Score 293.4; DB 1; Length 1622;
Best Local Similarity 77.4%; Pred. No. 4.3e-59;
Matches 400; Conservative 0; Mismatches 101; Indels 16; Gaps 3;

QY 1206 GGAAGCACAGGGGCCACTCGAGTCCGCCGAGAACCCACTGGAAGGGAACACCTGCCTCT 1265
DB 107 GGCAGGAGCTGAGCAGTGATGAATGGCTTACTGAGACACTGGATGGGAACGCCTGCTGG 166

QY 1266 CCAGATGACAGAGCTGGCTCGCTGAAGGTGCTGCACAA-----TGACTCCG 1313
DB 167 CCAGGTCCATGGCTGGCTCCACTGCAGTGCCTTTCATGGAGATCGCGATGCTGCACACTG 226

QY 1314 ACATCTCTTCCCTGTCAGTGGCTGGCTCTACAGCCAGCAGAGATGCCCCCTCGGAA 1373
DB 227 ATGTACTCTTCCCTGTCAGTGGAGTGGCTCTATGTGTCAGACAGATGCCCTGTGGAA 286

QY 1374 GCCTGGAGAACGGGACAGGACAGAGGACCAAGTTCTCCGGATCTCGAATTCGGTACA 1433
DB 287 GTCTGGAGAATGGAACAGGAAACAGAAACCATATAATCCAGAGCCTGGCTTCGGTACA 346

QY 1434 GTGTGAAGCCAGCTCTCCAGGCCACGGAAGTCTCTGAGCAGCGCTGTTA-CTTCTGCCT 1492
DB 347 G---GGAAGCCAGTTCTCCAGGCCAAGGAAGTCTCTGAGCAGCTTGCTACTCTTCTGCCT 403

QY 1493 CAGTGCACAGTCCCATGACAAATTAGTGAACCTGCGCCAGGCCACTGTGCCATGATGAACA 1552
DB 404 CAGTGCCTGAGTCCCATGACAGTCCATGAATGCGCCAAAGCCATCGTGCCCATGATGAACA 463

QY 1553 GGAAGGATGAGCTGGAGGAGGAGAACAGATCACTCGAAACCTGCTGACCGTGAGATGG 1612
DB 464 GAAAGATGAATTGGAAGAGAGATGGTCACTGCGGAATCTCTTGACCGTGAGATGG 523

QY 1613 AGCACTCAGCGCGCTCCGGCAAGAGTGGACACCTTGAAGAGGAAGTGGCTGAAACAGG 1672
DB 524 AGCACTCAGCAGCGCTCCGGCAAGAGTGGATGCCCTGAGGCGGAAGGTGACAGAACAGC 583

QY 1673 AGGAGCGGACGGCATGAGTCCAGGCGCTGGCCAG 1709
DB 584 AGGAGCGCATGCCAAAGGTCCAGGCACTGGCCAG 620

RESULT 12
US-10-450-763-214
; Sequence 214, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 214
; LENGTH: 454
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (220)...(387)
; OTHER INFORMATION: 58% homologous to Homo sapiens Partial sequence of the clone
; OTHER INFORMATION: 9 protein, accession number W42400, Smith-Waterman Score=142.
US-10-450-763-214

Query Match      14.4%; Score 274.6; DB 8; Length 454;

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:40:00 ; Search time 92.0732 Seconds
(without alignments)
2537.148 Million cell updates/sec

Title: US-09-155-676b-2

Perfect score: 3093

Sequence: 1 XTGPGXMGSGSNXDKRQF.....FLWTFHVCEPINCFLSLKKK 604

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3058	98.9	604	2 AAW42400	Partial s
2	1594.5	51.6	541	4 AAE00683	Aae00683 Human TNF
3	1537.5	49.7	784	4 AAE00685	Aae00685 Human TNF
4	1537.5	49.7	813	4 AAE00684	Aae00684 Human TNF
5	1196	38.7	375	4 AAB931962	Aab931962 Human pro
6	388	12.5	91	4 AEG00223	Abg00223 Novel hum
7	302	9.8	520	4 ABB61499	Abb61499 Drosophil
8	236	7.6	131	4 ABBG28013	Abbg28013 Novel hum
9	183.5	5.9	720	4 ABBG2508	Abbg2508 Drosophil
10	183	5.9	1019	4 ABB11928	Abb11928 Human BG3
11	175.5	5.7	473	6 ADA54368	Ada54368 Human pro
12	175.5	5.7	473	7 ADJ70685	Adj70685 Human hea
13	171	5.5	1006	4 ABG21178	Abg21178 Novel hum
14	160	5.2	2845	8 ADO08046	Ado08046 Mouse pol
15	154	5.0	388	4 ABG21174	Abg21174 Novel hum
16	151	4.9	658	7 ADE07881	Ade07881 Novel pro
17	151	4.9	704	2 AAY29517	Ray29517 Human lun
18	151	4.9	704	3 AAB44467	Aab44467 Human lun
19	151	4.9	704	4 AAE13809	Aae13809 Human lun
20	151	4.9	704	7 ADD66499	Add66499 Human lun
21	151	4.9	704	7 ADE87753	Ade87753 Human lun
22	149	4.8	641	8 ADR08868	Adr08868 Human pro
23	145	4.7	546	5 ABB90404	Abb90404 Human pol
24	144.5	4.7	342	7 ADD71112	Add71112 Human int
25	144	4.7	600	2 AAY29861	Aay29861 Human sec

26	143	4.6	183	4	ABG211175	Abg211175 Novel hum
27	139.5	4.5	298	7	ADE87009	Ade87009 Human pan
28	139.5	4.5	561	7	ADF78154	Adf78154 Human ext
29	139.5	4.5	597	7	ADF78163	Adf78163 Human ext
30	137	4.4	286	4	AAM25825	Aam25825 Human pro
31	137	4.4	286	4	AAM25950	Aam25950 Human pro
32	137	4.4	286	4	ABB11042	Abb11042 Human sec
33	136.5	4.4	369	5	ABB89669	Abb89669 Human pol
34	136.5	4.4	382	4	AAB92481	Aab92481 Human pro
35	136.5	4.4	434	3	AAB41968	Aab41968 Human ORF
36	136.5	4.4	2742	3	AAB23012	Aab23012 Human APC
37	136.5	4.4	2842	2	AAR63508	Aar63508 Adenonato
38	136.5	4.4	2842	5	ABG30988	Abg30988 Human APC
39	136.5	4.4	2843	2	AAR26052	Aar26052 APC Gene
40	136.5	4.4	2843	2	AAR58634	Aar58634 Adenonato
41	136.5	4.4	2843	2	AAW11922	Aaw11922 Adenonato
42	136.5	4.4	2843	2	AAW35392	Aaw35392 Human ade
43	136.5	4.4	2843	2	AAW38370	Aaw38370 Human ade
44	136.5	4.4	2843	2	AAW76140	Aaw76140 Human APC
45	136.5	4.4	2843	2	AAW76144	Aaw76144 Human APC

ALIGNMENTS

RESULT 1
AAW42400
ID AAW42400 standard; protein; 604 AA.
XX XX

AC AAW42400;

XX 15-APR-1998 (first entry)

DE DE Partial sequence of the clone 9 protein.

XX KW Human tumour necrosis factor receptor-associated factor 2; TRAF2;

KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;

KW intracellular signalling activity; acute hepatitis;

XX KW autoimmune-induced cell death.

OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 1 /note= "not specified"

FT Misc-difference 6 /note= "not specified"

FT Misc-difference 8 /note= "not specified"

FT Misc-difference 13 /note= "not specified"

FT Misc-difference 15 /note= "not specified"

FT Misc-difference 37 /note= "not specified"

FT Misc-difference 271 /note= "not specified"

FT Misc-difference 274 /note= "not specified"

FT Misc-difference 334 /note= "not specified"

FT Misc-difference 348 /note= "not specified"

FT Misc-difference 354 /note= "not specified"

FT Misc-difference 355 /note= "not specified"

FT Misc-difference 359 /note= "not specified"

FT Misc-difference 363 /note= "not specified"

FT Misc-difference 405 /note= "not specified"

FT Misc-difference 405 /note= "not specified"

Misc-difference 549 /note= "not specified"
FT
FT Misc-difference 569 /note= "not specified"
FT
FT Misc-difference 570 /note= "not specified"
FT
FT
XX WO9737016-A1.
XX
XX 09-OCT-1997.
XX
XX 01-APR-1997; 97WO-IL000117.
XX
XX 02-APR-1996; 96IL-00117800.
XX
XX 26-AUG-1996; 96IL-00119133.
XX
XX (YEDA) YEDA RES & DEV CO LTD.
XX
XX Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;
XX
XX WPI: 1997-503101/46.
XX
XX N-PSDB; AAV03323.
XX
XX DNA encoding tumour necrosis factor receptor-associated factor binding
PT molecule - used for modulation or mediation in cells of the activity of
PT NF-kB.
XX
XX Disclosure; Fig 3D; 127pp; English.
XX
XX The present sequence represents the sequence encoded by the 5' end of
CC clone 9, and is a TRAF2 binding protein. Clone 9 is a partial clone,
CC which lacks most of its 5' end of the coding DNA sequence. A cDNA library
CC prepared from B-cells was screened for proteins that associate with
CC TRAF2, and clone 9 DNA isolated. The clone 9 protein is capable of
CC binding to at least amino acids 222-501 of TRAF2. The TRAF-2 binding
CC proteins can be used for modulation or mediation in cells of the activity
CC of NF-kappaB or any other intracellular signalling activity modulated or
CC mediated by TRAF2. TRAF-binding proteins are especially used for
CC prevention or treatment of pathological conditions associated with NF-kB
CC induction, e.g. acute hepatitis, autoimmune-induced cell death, e.g.
CC death of the beta langerhans cells or the pancreas that results in
CC diabetes, the death of cells in graft rejection, the death of
CC oligodendrocytes in the brain in multiple sclerosis, and AIDS-inhibited T
CC cell suicide which causes proliferation of the AIDS virus and hence the
CC AIDS disease. The proteins are also useful for screening of ligands
CC capable of binding to a protein, which are useful for modulating cellular
CC activity modulated/mediated by TRAF2
XX
XX Sequence 604 AA;
XX
XX Query Match 98.9%; Score 3058; DB 2; Length 604;
XX Best Local Similarity 100.0%; Pred. No. 7e-272;
XX Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 2 TGPXGXWMSGXNKKRQFLERLLDAVKQCQIRFXGRKEIASDSRVTCLCAQFEAVL 61
XX 2 TGPXGXWMSGXNKKRQFLERLLDAVKQCQIRFXGRKEIASDSRVTCLCAQFEAVL 61
XX
XX 62 QHGLKRSRGLATAAIAKQAGFASKTETEPFVWYVYVKEVNLKHELORFVSLRHIAADVG 121
XX 62 QHGLKRSRGLATAAIAKQAGFASKTETEPFVWYVYVKEVNLKHELORFVSLRHIAADVG 121
XX
XX 122 RGRWLRCALNEHSLERYLHMLADRCRLSTFYEDWSFVWDEERSMPLPTMAAGLSILF 181
XX 122 RGRWLRCALNEHSLERYLHMLADRCRLSTFYEDWSFVWDEERSMPLPTMAAGLSILF 181
XX
XX 182 AINIDNKLNGQSKFAFTVSDLLKKESTQNTVSLTKKESTQVSSLFREITASSAVSLIKP 241
XX 182 AINIDNKLNGQSKFAFTVSDLLKKESTQNTVSLTKKESTQVSSLFREITASSAVSLIKP 241
XX
XX 242 EQETDPCLCSPGMSVLPNNAKSGRRRKKXPTXSHLMWRKMSRTLTGTLKRHLGQGAQR 301
XX 242 EQETDPCLCSPGMSVLPNNAKSGRRRKKXPTXSHLMWRKMSRTLTGTLKRHLGQGAQR 301
XX
XX

QY 302 TTPTAPLSISCPPLKAPSGLTPMSEQOLMENXFPVFERGVWVPEAXCEKHXRCXKX 361
Db 302 TTPTAPLSISCPPLKAPSGLTPMSEQOLMENXFPVFERGVWVPEAXCEKHXRCXKX 361
QY 362 RXRVWKLIRKEAQGPLGVAREATGREHLPLDPAQLGSAEGAAQXLRHPLPCQWRGLLOPS 421
Db 362 RXRVWKLIRKEAQGPLGVAREATGREHLPLDPAQLGSAEGAAQXLRHPLPCQWRGLLOPS 421
QY 422 RCPPRKPGERDRTPRSPFGSWTSVQCGSOLSRPRKSEQPVTSASVPESMTISELROAT 481
Db 422 RCPPRKPGERDRTPRSPFGSWTSVQCGSOLSRPRKSEQPVTSASVPESMTISELROAT 481
QY 482 VAMNRKDELEENRSLRNLLDGEHSAALRQEVDTLKRKVAEQEERQGMKVQALASYL 541
Db 482 VAMNRKDELEENRSLRNLLDGEHSAALRQEVDTLKRKVAEQEERQGMKVQALASYL 541
QY 542 CYFVRRFXPHVTRMWRNGREKSNSSXXSHLSWISQSLKCLFLWTFHVCEPINCFSHL 601
Db 542 CYFVRRFXPHVTRMWRNGREKSNSSXXSHLSWISQSLKCLFLWTFHVCEPINCFSHL 601
QY 602 KKK 604
Db 602 KKK 604
RESULT 2
AAE00683
ID AAE00683 standard; protein; 541 AA.
XX AC AAE00683;
XX DT 04-JUL-2001 (first entry)
XX DE Human TNF receptor-associated factor (TRAF2) binding protein, IREN.
XX KW Human; Tumour Necrosis Factor; TNF; TNF Receptor Associated Factor;
KW TRAF2; TRAF2 binding protein; IREN; IkappaB Regulator; immunosuppressive;
KW nuclear factor-kappaB; NF-kappaB; cytostatic; tumour; AIDS;
KW acquired immune deficiency syndrome; rheumatic disease; apoptosis;
KW autoimmune disease; septic shock; graft-vs-host reaction; inflammation;
KW anorexia; anti-HIV; therapy.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Binding-site 198..388
FT Region 326..331
FT /note= "This motif is also present within 51 amino acid
FT domain spanning amino acids 769-820 of NIK (Nuclear
FT factor (NF)-kappaB inducing kinase) which is essential
FT for IKK-1 (IkappaB kinase) binding to NIK in a 2-hybrid
FT assay and NF-kappaB activation by NIK over expression"
XX WO2000116314-A1.
XX
XX 08-MAR-2001.
XX 31-AUG-2000; 2000WO-IL0000517.
XX 02-SEP-1999; 95IL-00131719.
XX (YEDA) YEDA RES & DEV CO LTD.
XX Wallach D, Malinin N, Sinha I, Leu S;
XX WPI: 2001-281387/29.
XX N-PSDB; AAD04333.
XX New DNA sequence encoding Tumour Necrosis Factor receptor associated
FT factor (TRAF) binding proteins (IREN) for treatment or prevention of
PT pathological conditions associated with NF-kappaB induction.
PT

XX Claim 17; Fig 6; 118pp; English.

XX The present sequence is human tumour necrosis factor (TNF) receptor-associated factor (TRAF2) binding protein, IREN. A fragment of this IREN molecule is capable of binding to TRAF2 protein at position 225-501. The invention relates to human tumour necrosis factor (TNF) receptor-associated factor (TRAF2) binding protein designated as IREN (IkappaB Regulator), its isoforms IREN-10B, IREN-E and their corresponding cDNA molecules. IREN is useful for modulating/mediating the activity of intracellular signalling factor NF (Nuclear Factor)-kappaB or any other the prevention and treatment of a pathological condition associated with NF-kappaB induction (abnormal) e.g. AIDS (acquired immune deficiency syndrome), autoimmune diseases, tumours, rheumatic diseases, anorexia, septic shock and graft-vs-host reactions. IREN also plays an important role in the control of inflammation and other non-apoptotic effects of TNF as well as in the control of apoptosis. The invention also relates to method for screening, identifying and producing a molecule capable of modulating activities mediated by IREN. IREN antibodies are useful for the purification of new proteins from different sources, including cell extracts or transformed cell lines, in addition IREN can be used in diagnostic purposes for identifying disorders related to abnormal functioning of cellular effects mediated directly by TRAF proteins

XX Sequence 541 AA;

Query Match 51.6%; Score 1594.5; DB 4; Length 541;
Best Local Similarity 61.8%; Pred. No. 2.5e-137;
Matches 366; Conservative 22; Mismatches 101; Indels 103; Gaps 13;

Qy 9 MSGSXNDKROFLERLLDAVKQCIQIRFXGKEIASDSRVTCICAQFEAVLOHGLKRS 68
Db 1 MSGSNDKROFLERLLDAVKQCIQIRFGKKEIASDSRVTCICAQFEAVLOHGLKRS 60

Qy 69 RGLALTAATAIKQAAGPASKTEPFWYVYKVELNKHQLQRFYSLRHIASDVGRGAWLR 128
Db 61 RGLALTAATAIKQAAGPASKTEPFWYVYKVELNKHQLQRFYSLRHIASDVGRGAWLR 120

Qy 129 CALNEHSLERYLHMLADRCLSTFYEDWSFVMDERSSMLPTWAAGLNSILFPAINDNK 188
Db 121 CALNEHSLERYLHMLADRCLSTFYEDWSFVMDERSSMLPTWAAGLNSILFPAINDNK 180

Qy 189 DLNQSKFAPTVSDLLKESTQNTVSLKESYSGVSSLFREITASSAVSILIKPQETDPC 248
Db 181 DLNQSKFAPTVSDLLKESTQNTVSLKESYSGVSSLFREITASSAVSILIKPQETDP- 239

Qy 249 LSCFGMSVLMPN-----AKSGRRKRXKPYXSHLMKMKSRRTGLTCLKHLGQGAQR- 301
Db 240 -----LPVSRNVGSADAKCKKRRKKVTTNIIISFDDDEQNSGCVFKKTPGAGESSED 294

Qy 302 -----TPTAPLS-----ISCPPLKAPSGLTPME-----SEQQLMENX 334
Db 295 NSDRSSNVMSAFSPGPNSSQSSNSWKIDSLSLNGEFGYQKLDVSKIDDEDVDENE 354

Qy 335 FPFVFERGVVPEACEKXK-----XXRCCKXRRXRVKMLIRKEAQGLGVARE- 382
Db 355 DDVY-----GNSGRKRGHSESPKLEGNCTC-----LSQWHSWAPLKVHLND 398

Qy 383 -----ATGREHLPLPAQGLGSAGAAQXLRHPLPCQWGLLQPSRCPKPPKGERDRTG 436
Db 399 SDILFPVSGVGSFADAPLGSLENGTGPDHVLFP-----DPLGRYSVEA 443

Qy 437 PRSPGSWTSVQCSQLSPKRSQPTVSAPVPSMTISLRQATVAMNRKDELEENR 496
Db 444 -SSPGH-----GSPLS-----SLPSSAVPESMTISLRQATVAMNRKDELEENR 489

Qy 497 SLRNLGEMHSAAALRQEVDTLKRKVAEOERQGMKVQALASYLCYFVRRF 548
Db 490 SLRNLGEMHSAAALRQEVDTLKRKVAEOERQGMKVQALASYLCYFVRRF 541

RESULT 3

AAE00685
ID AAE00685 standard; protein; 784 AA.
XX AC AAE00685;
XX DT 04-JUL-2001 (first entry)
XX DE Human TNF receptor-associated factor binding protein, IREN-E.
XX KW Human; Tumour Necrosis Factor; TNF; TNF Receptor Associated Factor;
KW TRAF2; TRAF2 binding protein; IREN; IkappaB Regulator; IREN-E;
KW immunosuppressive; nuclear factor-kappaB; NF-kappaB; cycostatic; tumour;
KW AIDS; acquired immune deficiency syndrome; rheumatic disease; apoptosis;
KW autoimmune disease; septic shock; graft-vs-host reaction; inflammation;
KW anorexia; anti-HIV; therapy.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Binding-site 198..388
FT Region 326..331
FT /note= "TRAF2 binding region"
FT /note= "This motif is also present within 51 amino acid
FT domain spanning amino acids 769-820 of NIK (Nuclear
FT factor (NF)-kappaB Inducing Kinase) which is essential
FT for IKK-1 (IkappaB Kinase) binding to NIK in a 2-hybrid
FT assay and NF-kappaB activation by NIK over expression"
XX W0200116314-A1.
XX PN 08-MAR-2001.
XX PD 31-AUG-2000; 2000WO-IL0000517.
XX PF 02-SEP-1999; 99IL-00131719.
XX PR (YEDA) YEDA RES & DEV CO LTD.
XX PA Wallach D, Malinin N, Sinha I, Leu S;
XX WPI; 2001-281387/29.
XX DR N-PSDB; AAD04335.
XX PS New DNA sequence encoding Tumor Necrosis Factor receptor associated
PT factor (TRAF) binding proteins (IREN) for treatment or prevention of
PT pathological conditions associated with NF-kappaB induction.
XX Claim 14; Fig 8; 118pp; English.
XX The present sequence is human tumour necrosis factor (TNF) receptor-associated factor (TRAF2) binding protein isoform, IREN-E. A fragment of this IREN-E molecule is capable of binding to TRAF2 protein at position 225-501. The invention relates to human tumour necrosis factor (TNF) receptor-associated factor (TRAF2) binding protein designated as IREN (IkappaB Regulator), its isoforms IREN-10B, IREN-E and their corresponding cDNA molecules. IREN is useful for modulating/mediating the activity of transcription factor NF (Nuclear Factor)-kappaB or any other intracellular signalling activity mediated by TRAF2. IREN is useful in the prevention and treatment of a pathological condition associated with NF-kappaB induction (abnormal) e.g. AIDS (acquired immune deficiency syndrome), autoimmune diseases, tumours, rheumatic diseases, anorexia, septic shock and graft-vs-host reactions. IREN also plays an important role in the control of inflammation and other non-apoptotic effects of TNF as well as in the control of apoptosis. The invention also relates to method for screening, identifying and producing a molecule capable of modulating activities mediated by IREN. IREN antibodies are useful for the purification of new proteins from different sources, including cell extracts or transformed cell lines, in addition IREN can be used in diagnostic purposes for identifying disorders related to abnormal functioning of cellular effects mediated directly by TRAF proteins

XX Sequence 784 AA;

Mon Jun 13 13:27:10 2005

us-09-155-676b-2.rag

Query Match 49.7%; Score 1537.5; DB 4; Length 784;
Best Local Similarity 61.2%; Pred. No. 7.6e-132;
Matches 356; Conservative 22; Mismatches 101; Indels 103; Gaps 13;

QY 9 MSGSNXNDKROFLERLLDAVKQCOIRFXGKKEITASDSRVTCLCAQFEAVLQHLKRS 68
Db 1 MSGSQNDKROFLERLLDAVKQCOIRFXGKKEITASDSRVTCLCAQFEAVLQHLKRS 60

QY 69 RGLALTAATAAKQAAGFASKTETEPVFVYVYVKEVNLKHELQRFYSLRHIASDVGRGAWLR 128
Db 61 RGLALTAATAAKQAAGFASKTETEPVFVYVYVKEVNLKHELQRFYSLRHIASDVGRGAWLR 120

QY 129 CALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERSMLPTMAAGLNSILFAINIDNK 188
Db 121 CALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERSMLPTMAAGLNSILFAINIDNK 180

QY 189 DLNGQSKFAPTVDLLKSTQNTVTLKKESTQVSLFREITASSAVSILIKPEQETDPC 248
Db 181 DLNGQSKFAPTVDLLKSTQNTVTLKKESTQVSLFREITASSAVSILIKPEQETDPC 239

QY 249 LSCPGMSVLMFN-----AKRSGRKRKXPTXSHLMRMKMSRTLTCLKRLHGOGRAOR- 301
Db 240 -----LPVYSRVNSADACKCKKXKVTNIIISFDDEDEQNSGDVFKTTPGAGESSED 294

QY 302 -----TTPTAPLS-----LSCPLKAPSLGTPME-----SEQOLMENX 334
Db 295 NSDRSSVNIIMSAFESPGFNGSQSSNSWKIDSLNGEFGYQKLDVKSIDDEVDENE 354

QY 335 FPFVFERGVVPEAKCEKRR-----XXRCGXKRRXRVKWLIRKEAQPLGVARE- 382
Db 355 DDVY-----GNSGKRRKHSPEKPLEGNTC-----LSQMSWAPLKVHND 398

QY 383 -----ATGREHLPLDPAOLGSAAGAAQRLHPLPCWGLLQPSRCPKPKGERDRRTG 436
Db 399 SDILFPVSGVGSYPADAPLGLSLENGTGPEDHVL-----DFGLRYSVEA 443

QY 437 PRSPGWSVTSVQCSQISRPKSEQPVTSASVPESMTISELQATVAMNRRKDELEENR 496
Db 444 -SSPGH-----GSPLS-----SLPDSASVPESMTISELQATVAMNRRKDELEENR 489

QY 497 SLRNLLDGEMHSAALRQEVDTLKRKVAEQERQGMKVQALA 538
Db 490 SLRNLLDGEMHSAALRQEVDTLKRKVAEQERQGMKVQALA 531

RESULT 4
AAE00684
ID AAE00684 standard; protein; 813 AA.

XX AC AAE00684;
XX DT 04-JUL-2001 (first entry)
XX DE Human TNF receptor-associated factor binding protein, IREN-10B.

XX KW Human; Tumour Necrosis Factor; TNF; TNF Receptor Associated Factor;
KW TRAF2; TRAF2 binding protein; IkappaB Regulator; IREN-10B;
KW immunosuppressive; nuclear factor-kappaB; NF-kappaB; cytostatic; tumour;
KW AIDS; acquired immune deficiency syndrome; rheumatic disease; apoptosis;
KW autoimmune disease; septic shock; graft-vs-host reaction; inflammation;
KW anorexia; anti-HIV; therapy.

XX OS Homo sapiens.

XX FH Binding-site
FT 198. .388
FT /note= "TRAF2 binding region"
FT 326. .331
FT /note= "This motif is also present within 51 amino acid
FT domain spanning amino acids 769-820 of NIK (Nuclear
FT factor (NF)-kappaB Inducing Kinase) which is essential
FT for IKK-1 (IkappaB Kinase) binding to NIK in a 2-hybrid
FT assay and NF-kappaB activation by NIK over expression"

XX WC200116314-A1.
XX 08-MAR-2001.
XX 31-AUG-2000; 2000WO-IL000517.
XX 02-SEP-1999; 99IL-00131719.
XX (YEDA) YEDA RES & DEV CO LTD.
XX Wallach D, Malinin N, Sinha I, Leu S;
XX WPI; 2001-281387/29.
XX N-PSDB; AAD04334.

XX New DNA sequence encoding Tumor Necrosis Factor receptor associated
XX factor (TRAF) binding proteins (IREN) for treatment or prevention of
XX pathological conditions associated with NF-kappaB induction.

XX Claim 15; Fig 7; 118pp; English.

XX The present sequence is human tumour necrosis factor (TNF) receptor-
XX associated factor (TRAF2) binding protein isoform, IREN-10B. A fragment
XX of this IREN-10B molecule is capable of binding to TRAF2 protein at
XX position 225-501. The invention relates to human tumour necrosis factor
XX (TNF) receptor- associated factor (TRAF2) binding protein designated as
XX IREN (IkappaB Regulator), its isoforms IREN-10B, IREN-E and their
XX corresponding cDNA molecules. IREN is useful for modulating/mediating the
XX activity of transcription factor NF (Nuclear Factor)-kappaB or any other
XX intracellular signalling activity mediated by TRAF2. IREN is useful in
XX the prevention and treatment of a pathological condition associated with
XX NF-kappaB induction (abnormal) e.g. AIDS (acquired immune deficiency
XX syndrome), autoimmune diseases, tumours, rheumatic diseases, anorexia,
XX septic shock and graft-vs-host reactions. IREN also plays an important
XX role in the control of inflammation and other non-apoptotic effects of
XX TNF as well as in the control of apoptosis. The invention also relates to
XX method for screening, identifying and producing a molecule capable of
XX modulating activities mediated by IREN. IREN antibodies are useful for
XX the purification of new proteins from different sources, including cell
XX extracts or transformed cell lines, in addition IREN can be used in
XX diagnostic purposes for identifying disorders related to abnormal
XX functioning of cellular effects mediated directly by TRAF proteins

XX SQ Sequence 813 AA;

Query Match 49.7%; Score 1537.5; DB 4; Length 813;
Best Local Similarity 61.2%; Pred. No. 8e-132;
Matches 356; Conservative 22; Mismatches 101; Indels 103; Gaps 13;

QY 9 MSGSNXNDKROFLERLLDAVKQCOIRFXGKKEITASDSRVTCLCAQFEAVLQHLKRS 68
Db 1 MSGSQNDKROFLERLLDAVKQCOIRFXGKKEITASDSRVTCLCAQFEAVLQHLKRS 60

QY 69 RGLALTAATAAKQAAGFASKTETEPVFVYVYVKEVNLKHELQRFYSLRHIASDVGRGAWLR 128
Db 61 RGLALTAATAAKQAAGFASKTETEPVFVYVYVKEVNLKHELQRFYSLRHIASDVGRGAWLR 120

QY 129 CALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERSMLPTMAAGLNSILFAINIDNK 188
Db 121 CALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERSMLPTMAAGLNSILFAINIDNK 180

QY 189 DLNGQSKFAPTVDLLKSTQNTVTLKKESTQVSLFREITASSAVSILIKPEQETDPC 248
Db 181 DLNGQSKFAPTVDLLKSTQNTVTLKKESTQVSLFREITASSAVSILIKPEQETDPC 239

QY 249 LSCPGMSVLMFN-----AKRSGRKRKXPTXSHLMRMKMSRTLTCLKRLHGOGRAOR- 301
Db 240 -----LPVYSRVNSADACKCKKXKVTNIIISFDDEDEQNSGDVFKTTPGAGESSED 294

QY 302 -----TTPTAPLS-----LSCPLKAPSLGTPME-----SEQOLMENX 334
Db 295 NSDRSSVNIIMSAFESPGFNGSQSSNSWKIDSLNGEFGYQKLDVKSIDDEVDENE 354

Qy	335	FPVFERGVVVPAXCEKRR	-----XXRCGXKRRKRWKVLIRKEAQGPILGVARE- 382
		:	
Db	355	DDVY	-----GNSSGRKHGRHSESPKPLEGNTC-----LSQWHSWAPLKLVLHND 398
		:	
Qy	383	-----ATGREHLPLPLDQOLGSAEGAAQXLRHPLPCQWRGLLQPSRCPPRKPGERDRTRG 436	
		:	
Db	399	SDILFPVSGVGSYPADAPLGSLNGTGPEDHVLV	-----DPGLRYSVEA 443
		:	
Qy	437	PRSPGSWTSVQCQSOLSRPKSSEQPVTTSVSPESMTIISLRQATVAMNRRKDELEENR 496	
		:	
Db	444	-SSPGH-----GSPLS-----SLLPSASVPESMTIISLRQATVAMNRRKDELEENR 489	
		:	
Qy	497	SLRNLLDGEHMSAALRQEVDTLKRKVAEQEERQGMKVQALA 538	
		:	
Db	490	SLRNLLDGEHMSAALRQEVDTLKRKVAEQEERQGMKVQALA 531	
		:	
RESULT 5			
AAB93962			
ID AAB93962 standard; protein; 375 AA.			
XX		AAB93962;	
XX			
DT	26-JUN-2001	(first entry)	
XX			
XX		Human protein sequence SEQ ID NO:14017.	
XX		Human; primer; detection; diagnosis; antisense therapy; gene therapy.	
KW			
XX			
OS		Homo sapiens.	
XX			
PN		EP1074617-A2.	
XX			
PD		07-FEB-2001.	
XX			
PF	28-JUL-2000;	2000EP-00116126.	
XX			
PR	29-JUL-1999;	99JP-00248036.	
XX	27-AUG-1999;	99JP-00300253.	
PR	11-JAN-2000;	2000JP-00118776.	
PR	02-MAY-2000;	2000JP-00183767.	
PR	09-JUN-2000;	2000JP-00241899.	
XX			
PA	(HELI-)	HELIX RES INST.	
XX			
PI	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;		
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;		
XX			
DR	WPI; 2001-318749/34.		

CC	detection and/or diagnosis of the abnormality of the proteins encoded by
CC	the full-length cDNAs. The primers allow obtaining of the full-length
CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC	represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC	oligonucleotides, all of which are used in the exemplification of the
CC	present invention
XX	
SQ	Sequence 375 AA;
	Query Match 38.7%; Score 1196; DB 4; Length 375;
	Best Local Similarity 82.2%; Pred. No. 6.8e-101;
	Matches 244; Conservative 11; Mismatches 30; Indels 12; Gaps 2;
Qy	9 MSGSXNKKRQFLLELLDAVKQCQIRFXGRKEIASDSRSRVTCCLCAQFEAVLQHGLKRS 68
Db	1 MSGSQNDKQFLLERLLDAVKQCQIRFGGRKEIASDSRSRVTCCLCAQFEAVLQHGLKRS 60
Qy	69 RGLALTAAILKQAAGFASKTETEPVFWYKYEVNLKHQLORFYSLRHIAISDVGRGAWLR 128
Db	61 RGLALTAAILKQAAGFASKTETEPVFWYKYEVNLKHQLORFYSLRHIAISDVGRGAWLR 120
Qy	129 CALNEHSRLRYLHMLADRCRLSTFFVEDWSFVMDEERSMPLPTMAAGLSILFAINIDNK 188
Db	121 CALNEHSRLRYLHMLADRCRLSTFFVEDWSFVMDEERSMPLPTMAAGLSILFAINIDNK 180
Qy	189 DLNQSKFAPTVDLLKESQNVTSLLKESTQGVSILFREITASSAVSILIKPEQETDPC 248
Db	181 DLNQSKFAPTVDLLKESQNVTSLLKESTQGVSILFREITASSAVSILIKPEQETDP - 239
Qy	249 LSCPGMSVLMPN-----AKSRGRRKRKKPTXSHLMRMKMSRTLGTCLRKHLGGGRA 299
Db	: : : : : : : : : : : : : : : : :
Db	240 -----LPVSVRNVSADACKCKERRKKKVNTNIISFDDEEQNSGDVFVKKTPGAES 291
RESULT 6	
ABG00223	ID ABG00223 standard; protein; 91 AA.
XX	
AC	ABG00223;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	Novel human diagnostic protein #214.
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder.
XX	
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US008631.
XX	
PR	31-MAR-2000; 2000US-00540217.
XX	
PR	23-AUG-2000; 2000US-00649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Drmanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
XX	
DR	N-PSDB; AAS64410.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
XX	Claim 20; SEQ ID NO 30582; 103pp; English.
PS	
XX	

PT biodiversity.
 PS Claim 20; SEQ ID NO 58372; 103pp; English.
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 131 AA;
 Query Match 7.6%; Score 236; DB 4; Length 131;
 Best Local Similarity 57.6%; Pred. No. 3.1e-13;
 Matches 53; Conservative 5; Mismatches 16; Indels 18; Gaps 4;
 QY 392 PDAGLGSA-----EGAAQLRHPLPCQWRGLLPSCPPKPGERDTRGPRSPGS 442
 Db 46 PETEPASAYGPFSTETAGCAEAGNSQP--WLN-----KCPKPKGERDTRGPRSPRS 99
 QY 443 WTSVQCSQLSRPKSSEQPVT---SASVPES 471
 Db 100 WTSVQCSQLSRPKSSEQPVTFLCLSRVHDN 131
 RESULT 9
 ID ABB62508 standard; protein; 720 AA.
 AC ABB62508;
 XX
 XX 26-MAR-2002 (first entry)
 DT
 DE Drosophila melanogaster polypeptide SEQ ID NO 14316.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL06611.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 PS Disclosure; SEQ ID NO 14316; 21pp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL130511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABR57377-
 CC ABR72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 720 AA;
 Query Match 5.9%; Score 183.5; DB 4; Length 720;
 Best Local Similarity 21.8%; Pred. No. 2.4e-07;
 Matches 97; Conservative 61; Mismatches 162; Indels 125; Gaps 17;
 QY 14 NXDKRQFLERLLDAVKQCOIRFXGRKEIASDSRVTCLCAQFEAVLQHLKRSRGAL 73
 Db 28 NENAREIEFE-----VRQKSLEVLGVCE-----QTSALCTTLEALFLHGLKDS----F 71
 QY 74 TAAAIKQAGFASKTETEPVFTYVYKVLNKHQLQFYSLRHIAHSVDVGRGAWLRCAINE 133
 Db 72 LSATFNVIAGDVER-RPEPSFWAPCMVFMKQVIEQVGLSQITSETGQCRAWVROSLNE 130
 QY 134 HSLERYLHMLLADRCRLSTFYEDWSFVMDERSMPLTMAAGLNS-----ILFAINIDNKD 189
 Db 131 SVFSSYLNMRKNGSALSQYKRNALMRDSEG---LETAAKIMESLEAYVQFDLPVNSSL 187
 QY 190 LNCQSKFAPTVSDI---LK-----ESTQNTVTSLLKES 218
 Db 188 LNHWPDYSLQLSDLWTLPALKSCPISGVDVASSLGSDDIIAIPTPQLQTNELFSESISNS 247
 QY 219 TQGVSSLF--REITASSAVSILIKPEQETDFCLSCPGMSVLMPNKASGRRK-----268
 Db 248 PFCRNENFGVPFELSORENLDLLIQKDEMD-----VINEETGGDKAETNETAE 296
 QY 269 RKYPTXSHLMRMRKSRTLGCLKRLHGLGQGAQRTTTPAPLSISCPPLKAPSG-----320
 Db 297 QEAPSGSKTRSRKKS-----NRQMDSSSFADLDLEAPSL-LPQGSNSLSNLM 342
 QY 321 -----LTPMESEQQLMENXFPVFERGVWVP-----EAXCEKHXRXR 356
 Db 343 QTSWGGDLEATPTSPTEELTGSRF--FORSVSVSSVSLRSPSTTDRCSYNALLRKHSNR 400
 QY 357 CGKRRXRVRWKLIRKEAQGPGLVAR 381
 Db 401 ESGAGWSEIWEKPRASASN-LNVAQ 424
 RESULT 10
 ID ABB11928 standard; peptide; 1019 AA.
 XX
 AC ABB11928;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human BG33_7 protein homologue, SEQ ID NO:2298.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;

PT New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 XX which the gene is involved, or as target molecules for gene therapy.

PS Claim 14; SEQ ID NO 1936; 205pp; English.

CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA5710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.

XX Sequence 473 AA;

Query Match 5.7%; Score 175.5; DB 6; Length 473;
 Best Local Similarity 28.5%; Pred. No. 7.2e-07;
 Matches 75; Conservative 35; Mismatches 118; Indels 35; Gaps 12;
 QY 5 GXGMSGXNKKQFLLERLLDAVKOCQIRFXGRK-----EIASDSRVTCCLCAQF 57
 DB 14 GGGGGGKSLSARNAVER--RNLTIVC--RFSVKTLIDRSCFTIDSSPEFNNFAIL 70
 QY 58 EAVLQHLGR-----SRGLATAAAIK-QAAGFASKTETEPVFWYVKEVLNK---HELQR 109
 DB 71 EQILSHRLKEISQSCRWLHLQIPLOQVTFWG--YESPRFWDYIRVACKVQNCICS 128
 QY 110 FYSLRHIASDVGRGAWLRCAALNEHSLERYLHMLLADRCRLSTPYEDWMSFVMDERSM 169
 DB 129 IENNVSSSRKGRWIRVALMEKHLSEYISTALRDFKTRRFYEDGAILGEE--ANML 187
 QY 170 PTMAAGLNSILFALINDKNGSKFAPTVSDLLKSTQNTVSLK--ESTQGVSSILFR 227
 DB 188 AGMLLGLNAIDFSCFKGEGLDGS---FPAVID-----YTPYLYIOSSDSISSDEE 236
 QY 228 EI-TASSAVSILIKPEQETDPC 249
 DB 237 ELRTLGSGSSESTPENVGPPFL 259

RESULT 12
 ADJ70685
 ID ADJ70685 standard; protein; 473 AA.
 AC ADJ70685;
 XX
 XX 06-MAY-2004 (first entry)

DE Human heat mitochondrial protein as a therapeutic target SeqID2491.
 XX
 KW mitochondrial; human; screening assay; diabetes mellitus;
 KW Huntington's disease; osteoarthritis;
 KW Leber's hereditary optic neuropathy; LHON;
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
 KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;
 KW osteopathic; ophthalmological; cytostatic.

OS Homo sapiens.

XX WO2003087768-A2.

XX 23-OCT-2003.

XX 04-APR-2003; 2003WO-US010870.

XX 12-APR-2002; 2002US-0372843P.

XX 17-JUN-2002; 2002US-0389987P.

XX 20-SEP-2002; 2002US-0412418P.

XX (MITO-) MITOKOR.

XX (BUCK-) BUCK INST AGE RES.

XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;

PI Warnock DE;

XX WPI; 2003-845369/78.

XX Identifying a mitochondrial target for drug screening assays and for
 PT treating diseases associated with altered mitochondrial function,
 PT comprises detecting a modified polypeptide in a sample and correlating
 PT with the disease.

PS Claim 1; SEQ ID NO 2491; 180pp; English.

XX This invention relates to novel mitochondrial targets that can be used
 CC for therapeutic intervention in treating a disease associated with
 CC altered mitochondrial function. Specifically, it refers to a method for
 CC identifying proteins of the human heart mitochondrial proteome that are
 CC useful for drug screening assays, as well as therapeutic targets. The
 CC present invention describes a method for identifying such proteins that
 CC can be used in the treatment of various diseases associated with altered
 CC mitochondrial function including diabetes mellitus, Huntington's disease,
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 CC compositions have neuroprotective, nontropic, antidiabetic,
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 CC cytostatic activities. This polypeptide sequence is a human heart
 CC mitochondrial protein of the invention.

XX Sequence 473 AA;

Query Match 5.7%; Score 175.5; DB 7; Length 473;
 Best Local Similarity 28.5%; Pred. No. 7.2e-07;
 Matches 75; Conservative 35; Mismatches 118; Indels 35; Gaps 12;
 QY 5 GXGMSGXNKKQFLLERLLDAVKOCQIRFXGRK-----EIASDSRVTCCLCAQF 57
 DB 14 GGGGGGKSLSARNAVER--RNLTIVC--RFSVKTLIDRSCFTIDSSPEFNNFAIL 70
 QY 58 EAVLQHLGR-----SRGLATAAAIK-QAAGFASKTETEPVFWYVKEVLNK---HELQR 109
 DB 71 EQILSHRLKEISQSCRWLHLQIPLOQVTFWG--YESPRFWDYIRVACKVQNCICS 128
 QY 110 FYSLRHIASDVGRGAWLRCAALNEHSLERYLHMLLADRCRLSTPYEDWMSFVMDERSM 169
 DB 129 IENNVSSSRKGRWIRVALMEKHLSEYISTALRDFKTRRFYEDGAILGEE--ANML 187
 QY 170 PTMAAGLNSILFALINDKNGSKFAPTVSDLLKSTQNTVSLK--ESTQGVSSILFR 227
 DB 188 AGMLLGLNAIDFSCFKGEGLDGS---FPAVID-----YTPYLYIOSSDSISSDEE 236
 QY 228 EI-TASSAVSILIKPEQETDPC 249
 DB 237 ELRTLGSGSSESTPENVGPPFL 259

RESULT 13

ABG21178

ID ABG21178 standard; protein; 1006 AA.

XX AC ABG21178;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #21169.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

Qy 80 QAAGFASKT---ETEPVFWYVYKVLNKHLELQRFYSLRHIAADVGRCRAWLRCAALNEHSL 136
Db 2002 QASGYAPKSFHVEDTPVCF-----SRNSSLSLSID-----SEDDLL 2038
Qy 137 ERYLHMLADRCRLSTFYEDWSFVMDERSMLPTMAAGLSILFAINIDNKDL---NGQ 193
Db 2039 QECISSAMPKKRPSR-----LKSESEKQSPKVGGI--LAEDLTLDKDLQRPDSE 2088
Qy 194 SKFAPTVDL-----LKSTQNTVTLKKESTQGVSSLFREITASSAVSILK----- 240
Db 2089 HAFSPGSENFOWKAIQEGANSIVSLHQAASACLSQASSDSDLSLSKLSIGSLGSPF 2148
Qy 241 ---PEQETDCLSCPGKSVLMP-----NAX 262
Db 2149 HLTDPQBEKPTSNKGRILPKGBKSTLEAKKIESENKGIKGGKVKYKSLITGKIRNSE 2208
Qy 263 RSGRRKRXKPTXSHLMWRKMSRTLTCLKRLHGOGRQRTTPTAPLSICPPLKAPSGLT 322
Db 2209 ISSQWKQPLTNMPSISR--GRTM-----IHI-FGLRNSSSSTSPVKKGPPLKTPASKS 2260
Qy 323 PMESEQQLMENXFPVFERGVWVPEACEKHXXRCGXKRRXRVWVKLIRKEAQGPLGVARE 382
Db 2261 PSEG-----PGATTSPRGTPKAGKSELSPITRQTSQISGSKNGSSRS 2302
Qy 383 ATGREHLPPLPAQLGSAEGAAOXLRHPLPCOWRGLLOPSR---CPPRKPGRDRTRGP-- 437
Db 2303 GS-RDSTPSRPTQ-----QPLSRPMQSPGRNISPGRNGISFPNKLSQLPRTSPST 2353
Qy 438 ---RSPG---SWTSVQSGQLSPR--KSSEQPVTSASVPESMTISE-LRQATVAM-MN 486
Db 2354 ASTKSSSGKWSYISP--GROLSQNLTKQASLSKNASSIPRSSASKGLNQMGNGSN 2411
Qy 487 RKDELEENRSLRNLIDGE-MEHSAAALRQEV-----DTLKRKXVABQEE----- 528
Db 2412 KKVLSRMSSTKSSGSDSSEPALVRQSTFIKEAPSPTLRRKLEESASPESLSPSSRP 2471
Qy 529 ---RQMKVQALASYLCYFVRFXPHVTRTWWR-----NGGR-EKSNSSXXS 571
Db 2472 DSPTRSOATPVLSPLPDMSLSTHPSVOAGGWRKLPNLSPTTEYDNGRPTKRHDARS 2531
Qy 572 HLSS 575
Db 2532 HSES 2535
RESULT 15
ID ABG21174 standard; protein; 388 AA.
XX AC ABG21174;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #21165.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX FN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSEQ-) HYSEQ INC.
XX FI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.
DR N-PSDB; AAS85361.
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 51533; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 388 AA;
Query Match 5.0%; Score 154; DB 4; Length 388;
Best Local Similarity 25.7%; Pred. No. 5.1e-05;
Matches 53; Conservative 29; Mismatches 76; Indels 48; Gaps 6;
Qy 95 WY---YVKEVNLKHELFYSLRHIAADVGRCRAWLRCAALNEHSLRYLHMLADRCRLS 151
Db 53 WWRMDWTPRLPSRHHIIELEHLATFVNDVGRCAWLRLALNGLMECYLKLQEQARLR 112
Qy 152 TFYEDWSFVMDERSMLPTMAAGLSILFAINIDNKDLN-----GQSKFAPTVDLL 204
Db 113 EYQPTALLRDABEGEFLSLFQGLMSLSFELSYKSAILNEWTLTPLALSGLCP-LSELD 171
Qy 205 KESTQNTVTLKKESTQGV-----SSLFREITASSAVSIL 238
Db 172 PLSTSGAELQKESLDSISHSSESIEVHSHGHKTRRKNOKPTASSLSLOTASSQLSCS 231
Qy 239 IKPEQETDPCLCSPGMSVLMFNAKRS 264
Db 232 L-----NSDSCL-----LQENGSKS 246
Search completed: June 7, 2005, 11:57:19
Job time : 97.0732 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:43:10 ; Search time 20.8699 Seconds
(without alignments)
2784.626 Million cell updates/sec

Title: US-09-155-676B-2
Perfect score: 3093
Sequence: 1 XTPGXGXMGSGXNXXDKRQF.....FLWTFHVCEPINCFSLSLKKK 604

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score.	Query Match	Length	DB ID	Description
1	178.5	5.8	624	2 T27096	hypothetical prote
2	160	5.2	2845	2 I49505	adenomatous polypo
3	136.5	4.4	2843	1 RBHUP	adenomatous polypo
4	130.5	4.2	364	2 T43467	hypothetical prote
5	123	4.0	2272	2 C54972	voltage-dependent
6	122	3.9	2222	2 A37490	probable VPI/AB13
7	120.5	3.9	780	2 G84708	actin-interacting
8	118.5	3.8	1033	2 T37715	protein H3923.1
9	118	3.8	1246	2 G89287	probable serine/th
10	113.5	3.7	1192	2 T18611	centromere protein
11	113	3.7	905	2 A54654	microtubule-associ
12	112.5	3.6	5327	2 T13564	hypothetical prote
13	111.5	3.6	792	2 T29187	hypothetical prote
14	111	3.6	553	2 T45872	hypothetical prote
15	111	3.6	1291	2 T17242	beta spectrin, bet
16	111	3.6	2388	2 J80271	smooth muscle myos
17	110.5	3.6	1938	2 JC5421	protein M01A8.2
18	110.5	3.6	1972	2 JC5420	nuclear mitotic ap
19	109.5	3.5	893	2 G88551	hypothetical prote
20	109.5	3.5	2101	2 A42184	hypothetical prote
21	109	3.5	992	2 T46337	kinesin motor prot
22	109	3.5	1459	2 T30196	giantin - human
23	108.5	3.5	3225	2 I52300	Bassoon protein -
24	107.5	3.5	3942	2 T42730	hypothetical prote
25	106	3.4	623	2 T16167	homeotic protein C
26	106	3.4	975	2 S33121	basal transcriptio
27	106	3.4	1469	2 T09219	calcium channel pr
28	106	3.4	2223	2 A47447	microtubule-associ
29	106	3.4	2774	2 A43359	

30	105.5	3.4	699	2 T01029	hypothetical prote
31	105.5	3.4	840	1 S69204	pheromone response
32	105	3.4	1323	2 T00037	hypothetical prote
33	104.5	3.4	651	2 T15624	hypothetical prote
34	104.5	3.4	858	2 S15762	neurofilament trip
35	104	3.4	1008	2 S15762	hypothetical prote
36	104	3.4	1200	2 AE2304	neurofilament prot
37	103	3.3	1171	2 T00380	KIAA0637 protein -
38	103	3.3	3924	2 S37431	ankyrin 2, neurona
39	102.5	3.3	878	2 T17245	hypothetical prote
40	102	3.3	2441	2 S39161	CREB-binding prote
41	101.5	3.3	1332	1 I48314	homeotic protein C
42	101.5	3.3	3259	1 A56539	giantin - human
43	101	3.3	237	1 S47630	translation elonga
44	101	3.3	1616	2 I37183	gene APXL protein
45	101	3.3	2007	1 B43402	myosin heavy chain

ALIGNMENTS

RESULT 1

T27096

hypothetical protein Y51H1A.2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T27096

R;Smyle, R.

submitted to the EMBL Data Library, October 1998

A;Reference number: Z20309

A;Accession: T27096

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-624 <WIL>

A;Cross-references: UNIPROT:Q9XWJ7; EMBL:AL032644; PIDN:CAA21666.1; GSPDB:GN00020; CESP:

A;Experimental source: clone Y51H1A

C;Genetics:

A;Gene: CESP:Y51H1A.2

A;Map position: 2

A;Introns: 6/3; 57/2; 92/3; 151/2; 268/1; 414/3; 535/1

C;Superfamily: Caenorhabditis elegans hypothetical protein Y51H1A.2

Query Match	5.8%	Score 178.5	DB 2	Length 624	
Best Local Similarity	26.6%	Pred. No. 2.1e-05			
Matches	71	Conservative 44	Mismatches 117	Indels 35	Gaps 9
Qy	29	VKQCQTRFXGRKE-----IASDSRSVTC-----	LCAQFEAVLQGLKESRGLAL	73	
Db	24	LKQEKIRADLQKELDNVVVKAASPHSKRENVPSEITQNLCSNTEAIFHGLRDPFFKLG	83		
Qy	74	TAAAIKQAGFASKTETETEPVFWYVYKVLNKHQLQRFYSLRHIAASDVGRGRAWLRCAINE	133		
Db	84	TRYA-----KYPEPNFPVFSKFSHSRITSQIACLGQIRSEIGKSPAWIRIVLNE	133		
Qy	134	HSLEYVHMLLADRCRLSTFYEDWSP---VNDERSMPLPTMAAGLSILFAINIDNKDL	190		
Db	134	NALGQYQLDLAAEATAIQFYSDDAFLRLSDGQSERIRGLLKLPLSLPSIAATNSFL	193		
Qy	191	NGQSKAPTVDLLKSTQNTVSLKSTQVSLFREITASSAVSVLTKPEQETDPCLS	250		
Db	194	NTWTPPTLLIAGLMDGQPLKVGTLKARNPKPAHLTEEI-AIPDAIDLV-PEEDHD--IG	249		
Qy	251	CPGMSVLMPNAKRS-GRKKKKXPTXSH	276		
Db	250	SP--SYLEKRRRLRSRPIKRSNDDH	274		

RESULT 2

I49505
adenomatous polyposis coli protein - mouse
N;Alternate names: APC
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004

Db 2241 SSTSPVSKGKPLKTPASKSPSEGTATTS-----PRGA-KPSVKSE----- 2281

Qy 363 XRVWKLIRKEAQGLGVLGAREATGREHLPLPDAQLGSAEG-----AAQXLRHPLPCQWRGL 417

Db 2282 --LSPVARQTSQ--IGSSKAPSR-----GSRDSTPSRPAQQPLSRPIQSPGRNS 2328

Qy 418 LOPSR---CPRKPGERDRTGPRSPGWSVQCG--SOLSRPKSSEQPV----- 464

Db 2329 ISPRNGISPPNKLSQLPRTSSP--STASTKSSGGRKMSYTPGRCMSQQNLTKQTGLSKN 2387

Qy 465 SASYPESWTISE-LRQATVAM-MNRKDEL-----EENRSLRNLIDGEMEHSAAAL 512

Db 2388 ASSIPRESASAGLNQWNGANGANKVELSRMSSTKSSGSESDRSRPVL---VRQSTFI 2444

Qy 513 ROEYD-TLKRKVAE 525

Db 2445 KEAPSPTLRKLEE 2458

RESULT 4

T43467

hypothetical protein DKFZp434A1727.1 - human

C:Species: Homo sapiens (man)

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004

C:Accession: T43467

R:Bioecker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, December 1999

A:Reference number: 222513

A:Accession: T43467

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-364 <AAA>

A:Cross-references: UNIPROT:Q9UF50; EMBL:AL133657

A:Experimental source: adult testis; clone DKFZp434A1727

C:Genetics:

A>Note: DKFZp434A1727.1

C:Superfamily: human hypothetical protein DKFZp434A1727.1

Query Match 4.2%; Score 130.5; DB 2; Length 364;

Best Local Similarity 25.0%; Pred. No. 0.04;

Matches 53; Conservative 36; Mismatches 98; Indels 25; Gaps 7;

Qy 41 EIADSDSRVTCLCAQFEAVLQHLKRSRGLALPAAIKQAAGPASKTETEPVWYVKE 100

Db 49 EPIDDSSEEFVNFAAILEQLSHRFK-----ACAPAGPVSWFSS--DGQGFWDYIRL 99

Qy 101 VLNK---HELQRFYSLRHIAISDVGRGAWLRCALENEHSLERYLHMLADRCLSTFFYEDW 157

Db 100 ACSKVPNNCVSSIIENMENISTARAKGAWIRKVALMEKMSGYITLARDTRTTRRFYDSG 159

Qy 158 SFVNDSESSMLPTMAAGLSILPAINDKDLNGQSKFAPTVSDLLKESTQNVTSLLKE 217

Db 160 AIMLRDE-ATILTMGLSLAIDSFCLKGEVLGDKT---PVVID-----YTPYLK- 206

Qy 218 STQGVSSLFREITASSAVSILIKEQETDPCLSQPGMSVLMPNKAKSGRRKRK--- 270

Db 207 FTQSYDYLTDEERHSAESSTSDNSPEHPVL 238

RESULT 5

C54972

voltage-dependent calcium channel alpha 1E - mouse

C:Species: Mus musculus (house mouse)

C>Date: 12-Apr-1995 #sequence_revision 12-Apr-1995 #text_change 09-Jul-2004

C:Accession: C54972

R:Williams, M.E.; Marubio, L.M.; Deal, C.R.; Hans, M.; Brust, P.F.; Philipson, L.H.; Mil J. Biol. Chem. 269, 22347-22357, 1994

A>Title: Structure and functional characterization of neuronal alpha-1E calcium channel

A:Reference number: A54972; MUID:94350992; PMID:8071363

A:Accession: C54972

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra

A:Molecule type: mRNA

A:Residues: 1-2272 <WIL>

A:Cross-references: UNIPROT:Q61290; GB:L29346; NID:G522330; PIDN:AAA59206.1; PID:G522331

A>Note: authors translated the codon AGG for residue 788 as lys, and CCT for residue 886

C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 4.0%; Score 123; DB 2; Length 2272;

Best Local Similarity 20.8%; Pred. No. 1.6;

Matches 90; Conservative 58; Mismatches 181; Indels 104; Gaps 19;

Qy 155 EDWSFVNDSESSMLPTMAAGLSILPAINDKDLNGQSKFAPTVSDLLKESTQNVTSL 214

Db 658 EDWNEVMYNGIRSQ--GGVSSGMSAIYFIVL---TLFGNTYLLNVFLAIVDLNLANAQEL 713

Qy 215 LKSTQGVSSLFREITASSAVSILIKEQETDPCLSQPGMSVLMPNKAKSGRRKRK----- 270

Db 714 TKDE-QEIEEAFNQKHA-----LQKAKVSP-MSAPN---MPSIERDRRRHHMSMW 760

Qy 271 XPTXSHLMRKMGRSLTG-----TCLKRHLGQG-----RAQRTTPTAPLSICPPLKAPS 319

Db 761 EPRSLSLRERRRRHHMSVWEQRTSQLRRHMQMSQEAALNKEEAPPMPLN----PLNPLS 816

Qy 320 GLTPMSESQQLMENXFPVFERGVWVPEAKCEKHXXRCGXKRRXRVWKLIRKEAQGLGV 379

Db 817 PLNPLNAHPSLYRPRPIEGLAGLGLGKEEERISRG-----SLKGDIGG 863

Qy 380 ARSATGREHLPLPDAQLGSAEGAAQXLRHP--LPCQWRGLLQP-----SR 422

Db 864 LTSALDNQRSPL---SLGK-----REPPWLPKRSCHGNCDPQOBAGGETVVTFRDR 912

Qy 423 CPRKPGERDRTGPRSPGWSVQCSQLSRPKSSEQPVTSASVPESMTISBELRQATV 482

Db 913 ARHRQSQRRRHRVRTEG-----KDSASASRSASQ-----RSLDEGVSVGEKEHEP 963

Qy 483 AMWNRKDE--LEENR-----LRNLIDGEMEHSAAALRQBVDTLKRKVA 524

Db 964 HSSHRSKEPTIHEERTQDRLRTNSLMVPRGSLVGLDAETPLVQPOPLEYVKDAAL 1023

Qy 525 EQEERQGMKVOAL 537

Db 1024 TEQEAEGSSQAL 1036

RESULT 6

A37490

voltage-dependent calcium channel alpha 1E - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: A37490

R:Soong, T.W.; Stea, A.; Hodson, C.D.; Dubel, S.J.; Vincent, S.R.; Snutch, T.P. Science 260, 1133-1136, 1993

A>Title: Structure and functional expression of a member of the low voltage-activated ca

A:Reference number: A37490; MUID:93262464; PMID:8388125

A:Accession: A37490

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-2222 <SOO>

A:Cross-references: UNIPROT:Q07652; GB:L15453; NID:G310082; PIDN:AAA40855.1; PID:G310083

A:Experimental source: brain

A>Note: sequence extracted from NCBI backbone (NCBI:P:132101)

C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 3.9%; Score 122; DB 2; Length 2222;

Best Local Similarity 20.8%; Pred. No. 1.8;

Matches 90; Conservative 58; Mismatches 181; Indels 104; Gaps 19;

Qy 155 EDWSFVNDSESSMLPTMAAGLSILPAINDKDLNGQSKFAPTVSDLLKESTQNVTSL 214

Db 608 EDWNEVMYNGIRSQ--GGVSSGMSAIYFIVL---TLFGNTYLLNVFLAIVDLNLANAQEL 663

Qy 215 LKSTQGVSSLFREITASSAVSILIKEQETDPCLSQPGMSVLMPNKAKSGRRKRK----- 270

Db 664 TKDE-QEIEEAFNQKHA-----LQKAKVSP-MSAPN---MPSIERDRRRHHMSMW 710

Qy 271 XPTXSHLMRKMGRSLTG-----TCLKRHLGQG-----RAQRTTPTAPLSICPPLKAPS 319

Db 711 EPRSHLRERRHHMSVWEQRTSQLRRHMQMSQEAALNKEEAPPMPLN-----PLNPLS 766

Qy 320 GLTPMESEQOQWENKFFPVRGVVPEAXCEKHXRRXRCGXRRXRVWKLIRKEAQGLGV 379

Db 767 PLNPLNAPSLYRRPPIEGLALGLGLEKEEERISRG-----SLKGDIGG 813

Qy 380 AREATGRHPLPDAQLGSAEGAAQXLRHP--LPCWRGLLOPS-----R 422

Db 814 LTVSLVDNQRSL---SLGK-----REPPWLPSCRCHGNCDPQOETGGGETVVTPEDR 862

Qy 423 CPFRKPGERDRGRSPSGSWTSVQCQGLSRPRKSSQPVTSASVPESMTISLRQATV 482

Db 863 ARHROSQRSHRRVRTEKES-----ASASRSASOE-----RSLDEGVSDKEKEHEP 913

Qy 483 AMNRKDE--LEENRS-----LRNLDDGMEHSAALRQEVDTLKRKA 524

Db 914 QSHRSKPTIHEERTQDLRTNSLMVPRGSLVGLALDEAETPLVQPELEVGKDAAL 973

Qy 525 EBERQGMKVQAL 537

Db 974 TEQEAEGSSEQAL 986

RESULT 7

G84708

probable VP1/ABI3 family regulatory protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C:Accession: G84708

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, A Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; PMID:20083487; PMID:10617197

A;Accession: G84708

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-780 <STO>

A;Cross-references: UNIPROT:O04346; GB:AE002093; NID:g1946371; PIDN:AAB63089.1; GSPDB:GN

C;Genetics:

A;Gene: At2g30470

A;Map position: 2

Query Match 3.9%; Score 120.5; DB 2; Length 780;

Best Local Similarity 21.8%; Pred. No. 0.61;

Matches 122; Conservative 75; Mismatches 243; Indels 119; Gaps 30;

Qy 38 GRKEIASDSRVYTCACQAEAVLQHLKRSR-GLALTAALAKQAAGFASKTETE----- 91

Db 202 GMKHISPSQSNVHCASN---ILQ---KPSRPAISTPPVASKAQAARIGRPVVEGRG 255

Qy 92 ---PVFW---YYKVEV-----LNKHELQRPYSLRHIASDVR-GRWLRCALNEH----- 134

Db 256 HLLPRYPKTYDREVVQISGNLNLVPLPEKTL-ASDAGRIGRLVLPKACAEAYFPPI 314

Qy 135 SLERYLHMLAD-RCRLSTYEDWSFVW-----DEERSMLPTMAAGLSILF-----AIN 184

Db 315 SQSEGIPKIQDVRGR-----EWTFOFRYPWNNSRMVLEGVTPCIGSMMLQAGDVT 368

Qy 185 IDNKDLNGQ---SKFAPTVSDLLKESTQVNTSLKKESTQGVSSLPFREITASSAVSILIK 240

Db 369 FSRVDPGKILMSGRKKAAGDQGCGLTNGTSTEDTSSGVTVENPPSSGSCISLIPK 428

Qy 241 -----PEQETDPCLSGFMVMPNAKSGRRKRRKXPTXSHLMWRKMSRTLGLCKR--- 292

Db 429 ELANGPEN-----LNSETNGRIGDDPT--RVKCKRTRTIGAKNRLLL 471

Qy 293 HLGQGRAQRTTPTAPLISCPPLKAPSGLTPMESEQOLME-NXFPVF-ERGVMVPEAXCE 350

Db 472 HSESMELRLTWEEAQDLRLSPSVKPTIVVIE-EQEIEYDEPPVFGKTIYTKPSGE 530

Qy 351 KHXRCXGKRRXR-----VWKLIRKEAQGLGVARE-----ATGREH 388

Db 531 QERWATCDDCKWRRLPVDALLSFKWTCIDNVWDVSRCSAPESLKELENVLKVGREH 590

Qy 389 LPLPDAQLGSAEGAAQXLRHPPLPCWRGLLOPSRCPPR--KPGERDRTGRPSGWSVTSV 446

Db 591 ---KERTGCSQAQKQOE---PCGLDALASAAVLGDTTIGEPEVATTTTTPRHRACGSCI 644

Qy 447 QCGSQLSRPR-KSSEQPVTSASVPESMTISLRQATVAMNRKDELEENRSRLNLLDGE 505

Db 645 VC---IQPPSGKGRHKPTCGCTVCSTV---KRFKTLMRKRRKKQLERDVTAAEDKKKD 698

Qy 506 ME---HSAALRQEVDTLK 520

Db 699 MELAESDKSKEKEVNTAR 717

RESULT 8

T37715

actin-interacting homolog SPAC15A10.16 [imported] - fission yeast (Schizosaccharomyces P

C:Species: Schizosaccharomyces pombe

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000

C:Accession: T37715

R;Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, June 1997

A;Reference number: Z21738

A;Accession: T37715

A;Status: translated from GB/EMBL/DBD

A;Molecule type: DNA

A;Residues: 1-1033 <MUR>

A;Cross-references: EMBL:Z97208; PIDN:CA10112.1; GSPDB:GN00066; SPDB:SPAC15A10.16

A;Experimental source: strain 972h; cosmid c15A10

C;Genetics:

A;Gene: SPAC15E1.01; SPDB:SPAC15A10.16

A;Map position: 1

Query Match 3.8%; Score 118.5; DB 2; Length 1033;

Best Local Similarity 21.7%; Pred. No. 1.2;

Matches 95; Conservative 67; Mismatches 152; Indels 123; Gaps 23;

Qy 161 MDESSSMLPTMAAGLN-----SILFAINIDNK-----DLNGQSKFAPTVS 201

Db 177 LNQGRSN-LPOMTREANDNVNRNSRGSSQSGSISSFRNPDSPKQYEITTSASLSAFGAF 235

Qy 202 DLLKE-----STQNVTSLLK--ESTQGVSSLPFREITASSAVSILIKPEQETDPCLSCP 252

Db 236 DALQTNALERRASKRLSHVMVKDQNNEGSSQNLNITVES-VRGFLRDSKPDNIMDSP 294

Qy 253 G-----MSVLPNNAKSGRRKXKXPTXSHLMWRKMS--RTLGTCLKHLGQGRAQRTT 304

Db 295 SPKVSKRPSIVRODSHDSNKLRR-PTINTSFDKRFSPKLTRTSLTKSLDPG-----TP 348

Qy 305 TAPLSICPPL-KAPSGLTPMESEQOLMENFPVFERGVVPEAXCKHXXXRCXKRRX 363

Db 349 T---SLKSPSLKSPSSPVQKD-----VYSRS-----NSLRISOANRS 383

Qy 364 RVWKLIRKEAQGLGVAREATGREHLPLPDAQLGSAEGAAQXLRHPPLPCWRGLLOPSRC 423

Db 384 NVF-----PGATDVTRSVS--DHRILSSSTINDGEVAP-----PLP----- 417

Qy 424 PRKPGERDRT-RGRPSGWSVTSVQCQGLSRPR-KSSEQPV-----TSASVPES 471

Db 418 -----QSRSTISSPNSLATVLPSSPTILLPRGSRSTLSVNNKQFNADDGSLNSPNS 471

Qy 472 MTISLRQATVAMNRKDELEENRSRLNLLDGE--HSAALRQEVDTLKRKA----- 524

Db 472 IRETEEYAAAPKLEADIADDEVTDATSQRELLERQOKAESSEDTSEIRISLOGKLSLPQS 531

Qy 525 ---EQBERQGMKVQALAS 539

Db 532 STQOEIQFSSSVPEAAS 548

RESULT 9

G89287
A;Title: Genomic sequence of the nematode *C. elegans*: a platform for investigating biological protein H39E23.1 [imported] - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C;Accession: G89287
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genomic sequence of the nematode *C. elegans*: a platform for investigating biological protein H39E23.1 [imported]; PMID:9851916
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 287
A;Accession: G89287
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1246 <STO>
A;Cross-references: GB:chr_V; PIDN:CAB09532.1; PID:g3878100; GSPDB:GN00023; CESP:H39E23.1
C;Genetics:
A;Gene: H39E23.1
A;Map position: 5

Query Match 3.8%; Score 118; DB 2; Length 1246;
Best Local Similarity 19.9%; Pred. No. 1.7;
Matches 100; Conservative 75; Mismatches 211; Indels 116; Gaps 18;

Qy	128	RCALNEHSLERYLHMLLADRCRLSTFYEDMSFVMDER--SSMLPTMAAGLSILFAINID	186
Db	299	RSSLNIMKDRWNVGYEDD--ELKPFIEPPKQIDEQRIEKLIQIFOLGFNKAALLESVE	357
Qy	187	NKD-----LNGQSKFAPTVD-----LLKESTQNTVLSLKESTQGVSSLFRE--IT	230
Db	358	KEKPEDITHATYLLGERKSDMDASEITMAQSLLSHSSINVSSSLGQHPAGV--ITREHVT	415
Qy	231	ASSAVSILIKPEQETDPCSCPGMSVLMNPAKRSGRKRK-----XPTXSHLMRMKMSR	284
Db	416	SSSAGSSASPSRYRSRSATATGASITAGSALASANAQAQHQSSAAPSSGSSSSRRSQ	475
Qy	285	-----TLGTCLKRHLG--QGRAQRTTPTAPLSICPPLKAPSGLTPMSEQQLMEN	333
Db	476	NDAATAAGTVMGSTRHGGVQMRQPTSRQATISLLQPPSYKPSNTTQIAQIP----	531
Qy	334	XFVPFGRGVVPEAXCEKHXRCXGKRRXRVWKLIRKEAQGLGVAREATGREHLPLPD	393
Db	532	--PLFNRSNTATSSAAQ-----PSTGITGTRKIADPKGRIPLNS	568
Qy	394	AQL---GSAEGAAQXLRHPLPCQWRGLLQPSRCPPKPGERDRTRGPRSPGWSVTSVQGS	450
Db	569	TAVQGHRTATGAVANNGGI-----PSHRDHAQOQQYNNQLTSTTMS	611
Qy	451	QL-----SRPKSSEQPVTSASVPESMTISELRQATVAMNRRKDELE--EENRSLR	499
Db	612	KLINKTPAAGTAAATSSSSSSATSTAPLQSGSQISHAPTEPVIREDDDDNSENQNGN	671
Qy	500	NLLDGEHMSAALRQEVDTLTKRVAEERQGMKQALASVLYCFVRRFPVHRTVMWRN	559
Db	672	VPLIGGVGPQTSFAVQPTEDATSSDKEOQQQKASSETP-----KESNPV-----WQ	721
Qy	560	GGREKSNSSXXSHLSWIOQFL	581
Db	722	-----LHLSLTKSL	732

RESULT 10

T18611
N;Contains: probable serine/threonine-specific protein kinase (EC 2.7.1.-), long splice form - Caend
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18611; T18610; T23144; T23143
R;McMurray, A.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z18997
A;Accession: T18611

A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1192 <WILI>
A;Cross-references: UNIPROT:Q9TW45; EMBL:Z81027; PIDN:CAB54179.1; GSPDB:GN00023; CESP:H3
A;Experimental source: clone AH10
A;Accession: T18610
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-487,536-1192 <WIL2>
A;Cross-references: EMBL:Z81027; PIDN:CAB54178.1; GSPDB:GN00023; CESP:H39E23.1b
A;Experimental source: clone AH10
R;McMurray, A.
submitted to the EMBL Data Library, June 1997

A;Reference number: Z19696
A;Accession: T23144
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1192 <WIL3>
A;Cross-references: EMBL:Z96102; PIDN:CAB54263.1; GSPDB:GN00023; CESP:H39E23.1a
A;Experimental source: clone H39E23
A;Accession: T23143
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-487,536-1192 <WIL4>
A;Cross-references: EMBL:Z96102; PIDN:CAB54262.1; GSPDB:GN00023; CESP:H39E23.1b
A;Experimental source: clone H39E23
C;Genetics:
A;Gene: CESP:H39E23.1a; CESP:H39E23.1b
A;Map position: 5

A;Introns: 10/3; 33/3; 67/3; 139/2; 183/3; 451/3; 487/3; 535/3; 631/1; 825/2; 914/3; 992
C;Keywords: alternative splicing; ATP; phosphotransferase; serine/threonine-specific pro
F;1-1192/product: probable serine/threonine-specific protein kinase, long splice form #s
F;1-487,536-1192/product: probable serine/threonine-specific protein kinase, short splic

Query Match 3.7%; Score 113.5; DB 2; Length 1192;

Best Local Similarity 19.9%; Pred. No. 3.5;
Matches 91; Conservative 70; Mismatches 201; Indels 95; Gaps 15;

Qy	128	RCALNEHSLERYLHMLLADRCRLSTFYEDMSFVMDER--SSMLPTMAAGLSILFAINID	186
Db	409	RSSLNIMKDRWNVGYEDD--ELKPFIEPPKQIDEQRIEKLIQIFOLGFNKAALLESVE	467
Qy	187	NKD-----LNGQSKFAPTVD-----LLKESTQNTVLSLKESTQGVSSLFRE--IT	230
Db	468	KEKPEDITHATYLLGERKSDMDASEITMAQSLLSHSSINVSSSLGQHPAGV--ITREHVT	525
Qy	231	ASSAVSILIKPEQETDPCSCPGMSVLMNPAKRSGRKRK-----XPTXSHLMRMKMSR	284
Db	526	SSSAGSSASPSRYRSRSATATGASITAGSALASANAQAQHQSSAAPSSGSSSSRRSQ	585
Qy	285	-----TLGTCLKRHLG--QGRAQRTTPTAPLSICPPLKAPSGLTPMSEQQLMEN	333
Db	586	NDAATAAGTVMGSTRHGGVQMRQPTSRQATISLLQPPSYKPSNTTQIAQIP----	641
Qy	334	XFVPFGRGVVPEAXCEKHXRCXGKRRXRVWKLIRKEAQGLGVAREATGREHLPLPD	393
Db	642	--PLFNRSNTATSSAAQ-----PSTGITGTRKIADPKGRIPLNS	678
Qy	394	AQL---GSAEGAAQXLRHPLPCQWRGLLQPSRCPPKPGERDRTRGPRSPGWSVTSVQGS	450
Db	679	TAVQGHRTATGAVANNGGI-----PSHRDHAQOQQYNNQLTSTTMS	721
Qy	451	QL-----SRPKSSEQPVTSASVPESMTISELRQATVAMNRRKDELE--EENRSLR	499
Db	722	KLINKTPAAGTAAATSSSSSSATSTAPLQSGSQISHAPTEPVIREDDDDNSENQNGN	781
Qy	500	NLLDGEHMSAALRQEVDTLTKRVAEERQGMKQVA	536
Db	782	VPLIGGVGPQTSFAVQPTEDATSSDKEOQQQKASS	818

RESULT 11
A54654

Query Match	3.6%;	Score 112.5;	DB 2;	Length 5327;
Best Local Similarity	18.7%;	Pred. No. 29;		
Matches	80;	Conservative 64;	Mismatches 176;	Indels 107; Gaps 17;
QY	188	KDLNGSKFAPTVSDLLKESTQNVISL-----LKESTQGVSSLFRITAS	232	
	:	: :: :	:	:
Db	2936	KDQKAEBSRESVAESYKPESSKDATSAPPESKHSRPSVLGSLKD--EGDKTTSRRVSVA	2993	
	:	: :: :	:	:
QY	233	SAV-----SILI-----KPEQETDPC--LSCPGMSVLMPNAKSGRRKRKXPTXSHLMRK	281	
	:	: :: :	:	:
Db	2994	DSIKDEKSLVISOEASRPSAEBSLKDAAAPSOETSRPSVTSVDKGKSPVASKESRP	3053	
	:	: :: :	:	:
QY	282	MSRTLTGCTKRLHIGQGAQRITTPATLSISCPPLKAPS---GLTPM-----ESE	327	
	:	: :: :	:	:

[illegible][illegible]

Db 264 WYDFLKKLVYRHHGNNRPLLKIALTAVFGISIIIVAYRLIVKTRRRGGNIKKKPAK 323
Qy 240 -----KPEQETDPCLSCP--GMSVLMPNAKSGRRKRXPTX 274
Db 324 PLKRITPNNSDEKLKMKSPPKPLQATSDILQPLPSNNAWTPESKESANR---SPQS 380
Qy 275 SHL-MWRKMSLTGLTCLKRLHGGRA-----QRTTTPAPLSISCPPLKAPSGLTP 323
Db 381 SKINLVRNGGRVTNTDLKDCSEARAIPKCDKLTDFQTSTNSEQSPADAISAVAPTCLTG 440
Qy 324 MESQQLME-----NXPPVPERGVWP--EAXCEKHXKXRCGXKRRXRV--WKLI----- 369
Db 441 SOSSEQRQIEKMPNSFRELQDSFKFAEADIPPEERKSGSTEQSTTEPTKLAGSQSPS 500
Qy 370 -RKEAQGLGVAREATGREHLPLD--AQLGSABGAQXLRHPLPCQWRGLLQPSRCPPR 426
Db 501 ERQEIQQMPNSFRELQDSFKFAEADIPPEERKSGSTEQSTTEPTKLAGSQSPS 543
Qy 427 KPGERDRTRGPRSPGWSVQCGSQLRPRKSPQVTSASVPMSITSELURQATVAMN 486
Db 544 -----PCN-----NRRNSETSQVDTISPVPKLVGTQ-----SPS 573
Qy 487 RKDELEENRSLNLLDGEHSAALRQEVDTLKRKVAEQERQGMKVQALA 538
Db 574 ERQEIQQMPNSFRELQDSFKFAEADIPPEERKSGSTEQSTTEPTKLAGSQSPS 625

RESULT 14
T45872
hypothetical protein F4P12.20 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T45872
R;Blocher, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.
submitted to the Protein Sequence Database, January 2000
A;Reference number: 223016
A;Accession: T45872
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-553 <BL0>
A;Cross-references: UNIPROT:Q9LFI5; EMBL:AL132966
A;Experimental source: cultivar Columbia; BAC clone F4P12
C;Genetics:
A;Map position: 3
A;Introns: 35/1; 108/1; 197/1; 209/3; 231/3; 281/1; 491/3
A;Note: F4P12.20

Query Match 3.6%; Score 111; DB 2; Length 553;
Best Local Similarity 20.1%; Pred. No. 2;
Matches 93; Conservative 66; Mismatches 177; Indels 126; Gaps 22;
Qy 153 FYEDMSFVMDERSMLPTMAAGLSILFAINIDKOLNGSKFAPTVSDLLKESTQNTV 212
Db 103 FTSAGVLEPEELSSM-----ESNHKSGK-KALPTILEDINKSTESIS 145
Qy 213 SLKKESTQGVSS---LFREITA---SSAVSILIKPEQE-----TDPCLSCPGMSVLMPN 260
Db 146 TFQSDCTVENSQEFVLFDVTRASIQRSKTSVATPKSNVLRATDVAIS-PTSSTVDVT 204
Qy 261 AKRSGRRKRYPTYSHLMRMKMTGLTCLKRLHGGRAQRTTPTAPLSISCPPLKAPSG 320
Db 205 ATQGTKTSKSGP-----RNPSRV-----QCPGKATK-QPVATRGLSTISIKPPNG 248
Qy 321 LTPMESQQLMENFPFVPERGVWPVEAXCEKHXKXRCGXKRRXRVWKLIKEAQGP-LGV 379
Db 249 LSKVRPLSTSTN-----RSSLDISKQQEKNKSLPAG-----KEPLGPRISM 291
Qy 380 AREATGREHLPLDPAQLGSABGAQXLRHPLPCQWRGL---LQPSRCPPRKP----- 428
Db 292 SRR--KPVLPKGVFPKSSSRSDASKNEMTSSCSLSCASASSASHKPSIDSIKKK 349
Qy 429 -----GERDRTRG-----PRSPGWSVQCGSQLRPRKSPQVTSASVPESM 472

Db 350 NDSRSSLUSSQPLANRSTSGIMCQPRIPPOQT-----NKTSPKPLSSVSP-TAGSISDYS 403
Qy 473 TISELROATVAMNRRKDELEENRSLNLLDGEHSAALRQEVDTLKRKVAEQERQGM 532
Db 404 SESSRASETSKMANGNQKTVSREKVPAN--DNTVQTVKPLKNSKDT---SVQADAKEGT 458
Qy 533 K-VOALASYLCYFVRREXPVHRTMWRNGGREKSNSSXXSHL 573
Db 459 KRVSAL-----NGGLVPSASAKPSGL 479

RESULT 15

T17242

hypothetical protein DKFZp586B1417.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17242
R;Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18722
A;Accession: T17242
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1291 <KOE>
A;Cross-references: UNIPROT:Q9UFV0; EMBL:AL117448
A;Experimental source: adult uterus; clone DKFZp586B1417
C;Genetics:
A;Note: DKFZp586B1417.1

Query Match 3.6%; Score 111; DB 2; Length 1291;
Best Local Similarity 20.4%; Pred. No. 6;
Matches 39; Conservative 41; Mismatches 71; Indels 40; Gaps 7;
Qy 7 GXMGSXNXXDKRQFLERLLDAVK--QCQIRFXGRKEIASDSRVVTCICAQFEAVLQHG 64
Db 757 GLLEKCRNKTMRMLVEKMGREAVELGHGVNITGVEB-----NTLIASLCOLLERISHG 811
Qy 65 LKRSRGLA-----LTAALIKQAAGFASKTETE-----PVFWYVVK 99
Db 812 LQVKQKGSALWSHLLHYQDNQRKLTSGSL-STSGILLDSERRKSDASSLMPPL----- 864
Qy 100 EVLNKHELQRFYSLRHIASDVGRGAWLRCAALNEHSLELYLHMLADRCLSTFYEDWSF 159
Db 865 RISLIQDMRHIONIGEIKTDVGKARAWRLSMKLLSRHLKQLLSDHELTKKLYKRYAF 924
Qy 160 V-MDEERSMML 169
Db 925 LRCDDKEQFL 935

Search completed: June 7, 2005, 12:03:26

Job time : 24.2033 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:40:40 ; Search time 88.3902 Seconds
(without alignments)
3499.208 Million cell updates/sec

Title: US-09-155-676B-2
Perfect score: 3093
Sequence: 1 XTGPXGXMSGSXNXDKRQF.....FLWTFHVCBPINCQFSLKXK 604

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1196	38.7	375	2 Q9HA26	Q9ha26 homo sapien
2	400.5	12.9	249	2 Q8IU14	Q8iu14 homo sapien
3	358.5	11.6	373	2 Q8TE00	Q8teq0 homo sapien
4	329	10.6	476	2 Q9D353	Q9d3q3 mus musculus
5	308.5	10.0	249	2 Q8K050	Q8k050 mus musculus
6	302	9.8	520	2 Q9VKJ1	Q9vkj1 drosophila
7	257	8.3	1056	2 Q6P2R5	Q6p2r5 homo sapien
8	257	8.3	1058	2 Q9Y4G2	Q9y4g2 homo sapien
9	251	8.1	1074	2 Q7TS11	Q7ts11 mus musculus
10	190.5	6.2	1087	2 Q80TQ5	Q80tc5 mus musculus
11	189	6.1	1095	2 Q8IWE5	Q8iwe5 homo sapien
12	183.5	5.9	720	2 Q9W293	Q9w293 drosophila
13	183	5.9	1020	2 Q94928	Q94928 homo sapien
14	178.5	5.8	604	2 Q9XWJ7	Q9xwj7 caenorhabdi
15	175.5	5.7	473	2 Q96NLO	Q96n10 homo sapien
16	172	5.6	408	2 Q6PDC0	Q6pdc0 mus musculus
17	171	5.5	387	2 Q8TBG7	Q8tbg7 homo sapien
18	167	5.4	496	2 Q7QES7	Q7qes7 anopheles g
19	166	5.4	708	2 Q7L593	Q7l593 homo sapien
20	160	5.2	310	2 Q6P631	Q6p631 xenopus tro
21	160	5.2	712	2 Q8BIJ7	Q8bij7 mus musculus
22	160	5.2	2845	1 APC_MOUSE	Q6l315 mus musculus
23	159.5	5.2	1010	1 Y226_HUMAN	Q92622 homo sapien
24	157.5	5.1	961	2 Q6PD22	Q6pd22 mus musculus
25	155	5.0	329	2 Q8IWM5	Q8iwm5 homo sapien
26	151	4.9	600	1 RUF1_HUMAN	Q96t51 homo sapien
27	149	4.8	590	2 Q96PE1	Q96pe1 homo sapien
28	149	4.8	606	2 Q8WXA3	Q8wxal homo sapien
29	149	4.8	619	2 Q9P1Z1	Q9p1z1 homo sapien
30	148	4.8	600	2 Q9EPM6	Q9epm6 mus musculus
31	147.5	4.8	941	2 Q8BMP7	Q8bmp7 mus musculus

32	147.5	4.8	941	2 Q8BY22	Q8by22 mus musculus
33	146.5	4.7	606	2 Q8R4C2	Q8r4c2 mus musculus
34	145.5	4.7	628	2 Q69ZH1	Q69zh1 mus musculus
35	144.5	4.7	373	2 Q9CSV9	Q9csv9 mus musculus
36	143.5	4.6	941	2 Q6NMW8	Q6nmw8 mus musculus
37	142.5	4.6	556	2 Q8BL21	Q8bl21 mus musculus
38	141	4.6	469	2 Q68FK7	Q68fk7 xenopus lae
39	140.5	4.5	469	2 Q8VD10	Q8vd10 mus musculus
40	140.5	4.5	503	2 Q8BKQ4	Q8bkq4 mus musculus
41	139.5	4.5	413	2 Q8OY95	Q8oy95 mus musculus
42	139.5	4.5	446	2 Q08576	Q08576 mus musculus
43	139	4.5	428	2 Q7ZT23	Q7zt23 brachydanio
44	137.5	4.4	375	2 Q8N4U6	Q8n4u6 homo sapien
45	137.5	4.4	469	2 Q9D394	Q9d394 mus musculus

ALIGNMENTS

RESULT 1

Q9HA26	PRELIMINARY;	PRT;	375 AA.
AC Q9HA26;			
DT 01-MAR-2001 (Tremblrel. 16, Created)			
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)			
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)			
DE Hypothetical protein FLJ12363.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC TISSUE=Mammary gland;			
RX PubMed=14702039; DOI=10.1038/ng1285;			
RA Oca I., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,			
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,			
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,			
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,			
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,			
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,			
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,			
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,			
RA Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y., Ishida S.,			
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,			
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,			
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,			
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,			
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,			
RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,			
RA Nishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,			
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukushima Y.,			
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,			
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,			
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,			
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,			
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,			
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,			
RA Mizushima-Shigano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,			
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,			
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isoqai T., Sugano S.;			
"Complete sequencing and characterization of 21,243 full-length human			
cDNAs."			
RT Nat. Genet. 36:40-45(2004).			
RL EMBL; AK022425; BAB14033.1; --			
DR GO; GO:0003700; F:transcription factor activity; IEA.			
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.			
DR InterPro; IPR002197; HTH_Fis.			
DR InterPro; IPR004012; Run.			
DR Pfam; PF02759; RUN; 1.			
DR PRINTS; PRO1590; HTHFIS.			

```
DR SMART; SM00593; RUN: 1.
DR PROSITE; PS50826; RUN; 1.
SQ SEQUENCE 375 AA; 41861 MW; 7FD6BCBB4297CD95 CRC64;

Query Match 38.7%; Score 1196; DB 2; Length 375;
Best Local Similarity 82.2%; Pred. No. 1e-73;
Matches 244; Conservative 11; Mismatches 30; Indels 12; Gaps 2;

QY 9 MSGSNXDKRFLRLRLDAVKQOIRFXGKKEIASDSVTCCLCAQFEAVLOHGLKRS 68
Db 1 MSGSQNDKROFLRLRLDAVKQOIRFGGRKEIASDSVTCCLCAQFEAVLOHGLKRS 60

QY 69 RGLALTAARAIKQAGFASKTETEPVFWYVYVKEVLNKHLELQRFYSIRHIASDVGRGAWLR 128
Db 61 RGLALTAARAIKQAGFASKTETEPVFWYVYVKEVLNKHLELQRFYSIRHIASDVGRGAWLR 120

QY 129 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSMLPTWAAGLNSILFAINIDNK 188
Db 121 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSMLPTWAAGLNSILFAINIDNK 180

QY 189 DLNGOSKAPVTSVLLKKESTQNTVTSLLKKESTQGVSSLPREITASSAVSILIKPEQETDP 248
Db 181 DLNGOSKAPVTSVLLKKESTQNTVTSLLKKESTQGVSSLPREITASSAVSILIKPEQETDP- 239

QY 249 LSCPGMSVLPN-----AKRSGRKRXKPTXSHLMKMKMRTLTGTCCLKRHLGQORA 299
Db 240 -----LPVVSRRVNSADAKCKKRRKKKKVTNIIISFDDEDEQNSGDVFKKTPGAGES 291

RESULT 2
Q8IU14 PRELIMINARY; PRT; 249 AA.
AC Q8IU14;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE FLJ12363 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Primary B-Cells;
RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toehiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McWan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Primary B-Cells;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041583; AAH41583.1; -.
SQ SEQUENCE 249 AA; 27224 MW; 5F665B3F6BCE7FEB CRC64;
```

```
Query Match 12.9%; Score 400.5; DB 2; Length 249;
Best Local Similarity 64.8%; Pred. No. 1.7e-19;
Matches 94; Conservative 11; Mismatches 27; Indels 13; Gaps 4;

QY 161 MDERSMLPTWAAGLNSILFAINIDNKDINGOSKAPVTSVLLKKESTQNTVTSLLKKESTQ 220
Db 1 MDERSMLPTWAAAGNSILFAINIDNKDINGOSKAPVTSVLLKKESTQNTVTSLLKKESTQ 59

QY 221 GVSLSLREITASSAVSILIKPEQETDPCLSGPMVLPN-----AKRSGRKRXKPTXSH 276
Db 60 GVSLSLREITASSAVSILIKPEQETD-----LPVVSRRVNSADAKCKKRRKKKQVTNI 113

QY 277 LMM--RMSSTLTGTCCLKRHLGQORA 299
Db 114 ISFDDEDEQNSGDMFKKTPGAGES 138

RESULT 3
Q8TEQ0 PRELIMINARY; PRT; 373 AA.
ID Q8TEQ0;
AC Q8TEQ0;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE FLJ00143 protein (Fragment).
GN Name=FLJ00143;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074072; BAB84898.1; -.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001683; PX.
DR Pfam; PF00787; PX; 1.
DR SMART; SM00312; PX; 1.
DR PROSITE; PS50195; PX; 1.
FT NON_TER 1
SQ SEQUENCE 373 AA; 42541 MW; 2C34370BCCFABF7B CRC64;

Query Match 11.6%; Score 358.5; DB 2; Length 373;
Best Local Similarity 81.0%; Pred. No. 2.2e-16;
Matches 81; Conservative 1; Mismatches 5; Indels 13; Gaps 2;

QY 439 SPGSWTSVQCGSQLSRPRKSKSEQPVTSASPESMTISELROATVAMNRRKDELEENRSL 498
Db 5 SPGH-----GSLP-----SLTPSASVPESMTISELROATVAMNRRKDELEENRSL 51

QY 499 RNLLDGEHSAALRQEVDTLKRKVAEQEERQGMKQALA 538
Db 52 RNLLDGEHSAALRQEVDTLKRKVAEQEERQGMKQALA 91

RESULT 4
Q9D3S3 PRELIMINARY; PRT; 476 AA.
ID Q9D3S3;
AC Q9D3S3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
DE library, clone:4933437K13 product:hypothetical PX (Bem1/NCFI/P13K)
DE domain containing protein, full insert sequence (SNX29).
DE Name=4933437K13Rik; Synonyms=Snx29;
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
```



```

499 RNLLDGEHSAALRQEVDTLKRKVAQEERQGMKVQALA 538
|||||
50 RNLLDGEHSAALRQEVDALERRKVTQEQRHATKVQALA 89
|||||

RESULT 6
Q9VK31 PRELIMINARY; PRT; 520 AA.
AC Q9VK31.
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE CG5439-PA (LD35592p).
GN ORFNames=CG5439;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Betancur M.R., Bock J., Brokstein P., Brotter P.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Cantarel D.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fogle C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodgett J., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Taor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RA "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomic perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
[4]
RN SEQUENCE FROM N.A.
RP MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
[5]
RN SEQUENCE FROM N.A.
RP FLYBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
[6]
RN SEQUENCE FROM N.A.
RP FLYBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
[7]
RN SEQUENCE FROM N.A.
RP STRAIN-Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nuncio J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
[8]
RN SEQUENCE FROM N.A.
RP EMBL; AE003639; AAF3251.1; -
RA EMBL; AY095047; AAM11375.1; -
DR FLYBase; FBgn0032476; CG5439.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001683; PX.
DR InterPro; IPR004012; Run.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00787; PX; 1.
DR Pfam; PF02759; RUN; 1.
DR SMART; SM00312; PX; 1.
DR SMART; SM00593; RUN; 1.
DR PROSITE; PS00195; PX; 1.
DR PROSITE; PS00826; RUN; 1.
DR PROSITE; PS00678; WD_REPEATS; 1; UNKNOWN 1.
SQ SEQUENCE 520 AA; 59256 MW; 2BA4F6EB39AC605 CRC64;
Query Match 9.8%; Score 302; DB 2; Length 520;
Best Local Similarity 29.4%; Pred. No. 2.4e-12;
Matches 75; Conservative 59; Mismatches 99; Indels 22; Gaps 7;
QY 9 MSGSXXNXXKQFLLELLDAVKNQCOIRFXGRKEIASDSRVTCICAQFEVLQHLKRS 68
DB 22 LAGSTFGQRREDIFRRLQESAKHTISQKFSG-KELATERDESQVELCELEELMSYGLKRS 80
QY 69 RGLALTAATA-----IKQAAGFASKTETEPFVWYVYKVINKEHLQRFYSLRHASD 119
DB 81 AGTSFSFASFTQNNQWVSGNAGGS--NNDATWFCQTHLTHERQRYMDLQIWTN 139
QY 120 VGRGRAWLRCALNEHSLERYLHMLLADRCRLSTFTVEDWSFVMDERSMLPTMAAGLNSI 179
DB 140 VGRGRAFRATLNEKRLHSHVLTWLSDEQLHRFTWPSLLNDAAKKLPEIVDSLSDV 199
QY 180 LFAINDNKLNGQSKFAPTVSDLLKE-----STQNTVSLKESQTGVSSLFREITASSA 234

```

```

Db 200 L FALNVDITELNAPRRSTPSVA--VKBEPIFTTSPVPVVGQRKRPGL-AVERDIECVSS 256
Qy 235 VSIL---IKPQETD 246
Db 257 TEDLGLGALKPIESVE 271

RESULT 7
Q6P2R5 Q6P2R5 PRELIMINARY; PRT; 1056 AA.
AC Q6P2R5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Pleckstrin homology domain containing, family M (with RUN domain)
member 1.
GN Name=PLEKHM1;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J.J., Helton E., Kettman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2];
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strauberg R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; BC064361; AAH64361.1;
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR001849; PH.
DR InterPro; IPR004012; Run.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF02759; RUN; 1.
DR SMART; SM00109; C1; 1.
DR SMART; SM00233; PH; 2.
DR SMART; SM00593; RUN; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50826; RUN; 1.
SQ SEQUENCE 1056 AA; 117442 MW; 24004093CFA89079 CRC64;

Query Match 8.38; Score 257; DB 2; Length 1056;
Best Local Similarity 23.88; Pred. No. 7.3e-09;
Matches 122; Conservative 68; Mismatches 183; Indels 140; Gaps 21;

```

```

Qy 23 ERLIDAVKQCIIRPKGRKEITASDSRVTCLCAQFEAVLOHGLKRSGLTAATAAIKQAA 82
Db 20 KLVGVSKALQKQVSLDTVTYVSDGDANTMCSALEAVFIHGLHAKHIRAEAGGRKKS 79
Qy 83 GFASKTETEPVFWYVKEVLKHELQRFYSIRHTASDVGRGRAWLRKALNEHSLEYLHM 142
Db 80 --HQKPLPQVPFWPLLKAVTHKHIISELEHLTFVNTDVGRCRAWLRALNDGLMECYLKL 137
Qy 143 LLADRCRLSTFYEDWMSFVMDERSMMLPTMAAGLNSILFAINIDNKLN-----GQSK 195
Db 138 LLQEQARLHEVYQPTALLRDAEEGEPLSLFQGLTSLSFELSYKSAILNEWTLPLALSG 197
Qy 196 PAPTVDLLKESTQNTVSLKKESTQGVSSLF-----REITASS----- 233
Db 198 LCP-LSLDELSTSGAELQKESLDISHSSEDIIEVHSHGKHRRNQKLTASSLSDT 256
Qy 234 AVSILIKPEQETDPC-----SCPGMSVLMNAPKRRKRRKXPTXSHL 277
Db 257 ASSQSLSCSLNSCLQENGSKSPDHCPEPMSCDS-DLGTANAEDSDRSLOE----- 308
Qy 278 MMRKMSLTGLTKRHLGQGAQRT--TTPAPLS-----ISCPPLKAPSGLTPME 325
Db 309 VLLEFSKAQVNSVPTN---GLSQETEPTPQASLSLHGLNTSTYLHC---EAPAEPLPAQ 362
Qy 326 SEQOLMENXFPVFERGVWPEAXCEKHXRCGKRRXRVWKLIRKEAQGPLGVAR--EA 383
Db 363 AASGTQD-----GVHVQEP-----RQAPSPDLQQPVES 392
Qy 384 TGREHLPLDPA-----QLGSAEAGRAQLRHLPLPCOMGLQLQSPRCPKRGDRTRG 436
Db 393 TSGQQ---PSTVSETAREVGQNGLOKAQHD-----GAGLKLVVSS 432
Qy 437 PRSP--GSWTSVQCGSQLSRKSPKSEQPVTSAS 467
Db 433 PTSPKKNKSWIS---EDDFYRP--SREQPLESAS 460

RESULT 8
Q9Y4G2 Q9Y4G2 PRELIMINARY; PRT; 1058 AA.
AC Q9Y4G2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-MAR-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE KIAA0356 protein (Fragment).
GN Name=KIAA0356;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Nagase T., Kikuno R., Yamakawa H., Ohara O.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AB002354; BAA20813.2;
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR001849; PH.
DR InterPro; IPR004012; Run.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF02759; RUN; 1.
DR SMART; SM00109; C1; 1.
DR SMART; SM00233; PH; 2.
DR SMART; SM00593; RUN; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50826; RUN; 1.
FT NON TER 1
SQ SEQUENCE 1058 AA; 117670 MW; CDIAD1869C8C9D1 CRC64;

```

```

Query Match      8.3%; Score 257; DB 2; Length 1058;
Best Local Similarity 23.8%; Pred. No. 7.3e-09;
Matches 122; Conservative 68; Mismatches 183; Indels 140; Gaps 21;

QY 23 ERLDAVQCOIRFXGRKEIASDSRVTCCLCAQFEAVLQHLKRSGLALTAAAIKQAA 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 22 KKLGVSRALQKQVSLTVTSEDGDANTMCSEALVFIHGLHAKHRAEAGGKRKKA 81

QY 83 GFASKTEPEFVYVYVKEVNLKHELOFYSLRHIASDVGRGAWLRCAALNEHSLEYLHM 142
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 82 --HQKPLQFPWFVPLKATVHKHIISELEHTFTVNTDVGRCRAWLRALNDGLMECYKL 139

QY 143 LIADRCRLSTFYEDWSFVMDERSMLPTMAAGNSILFAINDKOLN-----GOSK 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 140 LIQEQARLHEYQPTALLRDAEAGEBFLSLFLQGLTSLSFELSYKSAIINWTLPLALSG 199

QY 196 FAPTVDLLKSTQNTVSLKKESTQGVSSLF-----REITASS----- 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 200 LCP-LSELPLSTSGAELQKESLDSHSSGSEDIHVHGHKIRRNQKLTAASLSLDT 258

QY 234 AVSILIKPEQETDPCL-----SCPGMSVLMNPNAKSGRRKRRKXPTXSHL 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 259 ASSQSLSCSLNSDSCLLQENGSKSPDHCEPMSCDSD-LGTANAEDSDSLQE----- 310

QY 278 MMRKMSRTLTGTLKRHLQOGRART---TPTAPLS-----ISCPLKAPSGLTPME 325
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 311 VLLEFSKAQVNSVPTN---GLSQETEIPTQASLSLHGLNTSTYLHC---BAPAEPLPAQ 364

QY 326 SEQQLMENXFPVPERGVVPEAXCEKHXRCXKRRXRVWKLIRKEAOGPLGVAR--EA 383
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 365 AASGTQD-----GVHVQEP-----RPQAPSLDLQOPVES 394

QY 384 TGREHLPLPDA-----OLGSAEGAAOXLRHPLPCWGRGLLQPSRCPKXPERDRTRG 436
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 395 TSCQQ---PSSIVSETAREVGGNGI-LQKAQHD-----GAGLKLIVSS 434

QY 437 PRSP--GSWTSVQCGSQLSRPKSSQPVTSAS 467
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 435 PTPSKNKSMTS---EDDFYRP--SREQPLESAS 462

RESULT 9
Q7TSII PRELIMINARY; PRT; 1074 AA.
AC Q7TSII
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Pleckstrin homology domain containing, family M (With RUN domain)
DE member 1.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Casavant T.L., Scheetz T.E.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Rubenstein P., Prange C.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Mullahy S.J.,
RA Raha S.S., Ioquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs K.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

```


QY 3 GPGXGXMSGX--NXDKRQFLERLLDAVKQCQIRFXG-RKEIAS--DSDSRVTCLCAQF 57

[illegible]

RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguichi S., Itoh T., Shigetani K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yanashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.",
RL Nat. Genet. 36:40-45 (2004).
DR EMBL; AK055233; BAB70882.1; -.
DR InterPro; IPR004012; Run.
DR Pfam; PF02759; RUN; 1.
DR SMART; SM00593; RUN; 1.
DR PROSITE; PS00826; RUN; 1.
SQ SEQUENCE 473 AA; 52813 MW; 1AC96C90379EEFB8 CRC64;
Query Match 5.7%; Score 175.5; DB 2; Length 473;
Best Local Similarity 28.5%; Pred. No. 0.00099;
Matches 75; Conservative 35; Mismatches 118; Indels 35; Gaps 12;
QY 5 GXGMSGSXNKKRQFILLERLLDAVQCQIRFXGRK-----ETASDSDSRVTCGLCAQF 57
Db 14 GGGGGGKKLSARNAVER-RNLIITVC--RFSVKTLIDRSCPETIDDSPEFNFAAIL 70
QY 58 EAVLQHGKLR-----SRGLATAAAIK-QAGFAASKTETEFVFWYVYKEVLNK---HELQR 109
Db 71 EQILSHRLKETISOSCRWLAHLQIPLOGQVTFG--YESPRFWDYIRVACRKVSQNCICS 128
QY 110 FYSLEHTASDVGRGMRCLALNEHSLRYLHMLLADRCRLSTFYEDWSFVMDDEERSMML 169
Db 129 IENNVSSSPRAKRAWRIRVALMEKHLSEYISTALRDFKTRFRFYEDGAIVLGE-ANML 187
QY 170 PTMAAGLNSILFALINDKDLNGSKFAPVSDLLKESTQNTVSLIK--ESTQGVSLFR 227
Db 168 AGMLLGLNAIDFSCLKGEGLDGS---FPAVID-----YTPYLKYOSSDSISDDEE 236
QY 228 EI-TASSAVSILIKPEQETDPCL 249
Db 237 ELRTLGSGSSESTPENVGPPFL 259
Search completed: June 7, 2005, 12:02:13
Job time : 92.3902 secs

RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguichi S., Itoh T., Shigetani K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yanashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.",
RL Nat. Genet. 36:40-45 (2004).
DR EMBL; AK055233; BAB70882.1; -.
DR InterPro; IPR004012; Run.
DR Pfam; PF02759; RUN; 1.
DR SMART; SM00593; RUN; 1.
DR PROSITE; PS00826; RUN; 1.
SQ SEQUENCE 604 AA; 66549 MW; AB317F77P2B63201 CRC64;
Query Match 5.8%; Score 178.5; DB 2; Length 604;
Best Local Similarity 26.6%; Pred. No. 0.00085;
Matches 71; Conservative 44; Mismatches 117; Indels 35; Gaps 9;
QY 29 VKOCQIRFXGRK-----IASDSDSRVTC-----LCAQFEAVLQHGKLSRGLAL 73
Db 24 LKQEKIRADLQKELDNVVAATASHSKSENVPSBITQNLNSIEAIFIHGLRDPFFLKG 83
QY 74 TAAAIKQAGAFASKTETEFVFWYVYKEVLNKHLELQRFYSLRHIAADVGRGRAWLRCALNE 133
Db 84 TRYA-----KYPEPFWFVSKFSHRGITSQIACLGQIRSEIGKSRAMIRIVLNE 133
QY 134 HSLERYLHMLLADRCRLSTFYEDWSF---VMDERSMSLPTMAAGLNSILFALINDKDL 190
Db 134 NALGOVLDLAAEATAIQOYSDAFRLRLSDGQSERIRGLKPLSLPISAATNSFL 193
QY 191 NGOSKFAPTVDLLKESTQNTVSLKKESTQGVSSLFREITASSAVSILIKPQETDPCLS 250
Db 194 NTWPTPLILAGLMDGQPLKVGTLKARNPKPAHLTEEI-AIPDAIDLV-PEEDHD--IG 249
QY 251 CPGMSVLMPNAKRS-GRKKKXPTXSH 276
Db 250 SP--SYLEKRRRRALSPIRKSENDH 274
RESULT 15
Q96NLO PRELIMINARY; PRT; 473 AA.
AC Q96NLO;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ30671.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX Pubmed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoa K., Chiba T., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Ishida S.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togliya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshihara Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:51:20 ; Search time 23.6321 Seconds
(without alignments)
1907.914 Million cell updates/sec

Title: US-09-155-676B-2

Perfect score: 3093

Sequence: 1 XTGPXKXMGSGXNXXDKRQF.....FLWTFHVCEPINCPSHLKXK 604

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/6C_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/6D_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	285.5	9.6	270	4	US-09-270-767-43983
2	151	4.9	704	4	US-09-370-838-191
3	151	4.9	704	4	US-09-854-133-191
4	136.5	4.4	2842	1	US-07-741-940-7
5	136.5	4.4	2842	1	US-08-289-548A-7
6	136.5	4.4	2842	1	US-08-452-654-7
7	136.5	4.4	2842	1	US-08-449-731-7
8	136.5	4.4	2843	1	US-07-741-940-2
9	136.5	4.4	2843	1	US-08-289-548A-2
10	136.5	4.4	2843	1	US-08-452-654-2
11	136.5	4.4	2843	1	US-08-452-655B-7
12	136.5	4.4	2843	1	US-08-452-655B-7
13	136.5	4.4	2843	2	US-08-370-235A-2
14	136.5	4.4	2843	3	US-08-450-582-2
15	136.5	4.4	2843	3	US-08-450-582-7
16	136.5	4.4	2843	4	US-08-449-731-2
17	136.5	4.4	2843	4	US-10-092-138A-30
18	136.5	4.4	2843	4	US-09-538-092-1007
19	136.5	4.4	2973	2	US-08-821-355A-7
20	136.5	4.4	2973	2	US-09-003-687A-7
21	136.5	4.4	2973	3	US-09-136-605-7
22	136	4.4	1139	4	US-09-949-016-6556
23	136	4.4	1146	4	US-09-949-016-7567
24	130.5	4.2	440	4	US-09-949-016-9774
25	128	4.1	601	4	US-09-252-991A-27821
26	120	3.9	427	4	US-09-252-991A-22649
27	117	3.8	329	4	US-09-252-991A-19752

28 115.5 3.7 458 3 US-09-352-159-9 Sequence 9, Appli
29 115.5 3.7 458 3 US-09-352-168-9 Sequence 9, Appli
30 115.5 3.7 458 4 US-09-771-045B-9 Sequence 9, Appli
31 115.5 3.7 458 4 US-09-770-564A-9 Sequence 9, Appli
32 113 3.7 432 4 US-09-252-991A-30848 Sequence 30848, A
33 112 3.6 622 4 US-09-252-991A-32308 Sequence 32308, A
34 112 3.6 2388 4 US-09-695-795A-2 Sequence 2, Appli
35 110.5 3.6 1972 4 US-08-875-435B-3 Sequence 3, Appli
36 110 3.6 326 4 US-09-252-991A-30438 Sequence 30438, A
37 109.5 3.5 2101 1 US-08-456-390-4 Sequence 4, Appli
38 109.5 3.5 2101 1 US-08-470-950-4 Sequence 4, Appli
39 108.5 3.5 2101 1 US-08-467-781-4 Sequence 4, Appli
40 109.5 3.5 2101 1 US-08-195-487-4 Sequence 4, Appli
41 109.5 3.5 2101 2 US-08-483-924-4 Sequence 4, Appli
42 109.5 3.5 2101 3 US-09-452-294-1 Sequence 1, Appli
43 109.5 3.5 2101 5 PCT-US93-06160-4 Sequence 4, Appli
44 108.5 3.5 2733 4 US-09-949-016-11433 Sequence 11433, A
45 108 3.5 541 4 US-09-252-991A-17206 Sequence 17206, A

ALIGNMENTS

RESULT 1

US-09-270-767-43983

; Sequence 43983, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 43983

; LENGTH: 270

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-43983

Query Match 9.6%; Score 295.5; DB 4; Length 270;

Best Local Similarity 31.7%; Pred. No. 2.7e-21;

Matches 64; Conservative 46; Mismatches 81; Indels 11; Gaps 3;

Qy 9 MSGSXNKKQFLLERLLDAVKQCQIRFXGRKEIASDSRVTCCLCAQFEAVLQHLKRS 68

Db 48 LAGSTFGQRAQDIFRRLQESAHKISQKFSG-KELATERDESQVCLCESLELMSYGLRQS 106

Qy 69 RGLALTA---IKQAAGFASKTETEPVFWYVYKVLNKHSLQRYSLRHASD 119

Db 107 AGTSSFSAAFIQNMQEMSGNAGGGS-NNNDATFWFCQTHLTPFRQRMDLQIWTN 165

Qy 120 VGRGRAWLRCALEHNSLYLHMLLADRCRLSFYEDWSFVMDERSMPLTMAAGNSI 179

Db 166 VGRGAFIRATLEKRLSHVLTWLSDEQLHRFYTPWLSLLNDEAKKLPFIVDSLSDV 225

Qy 180 LFAINDNKDLNGSKFAPTVS 201

Db 226 LFAINDVDTLNPARRSTPSVA 247

RESULT 2

US-09-370-838-191

; Sequence 191, Application US/09370838

; Patent No. 6444425

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Lodes, Michael J.

; APPLICANT: Mohamath, Roadoh

; APPLICANT: Secrist, Heather

Mon Jun 13 13:27:10 2005

```

; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 191
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Homo sapien
; ORGANISM: Homo sapien
US-09-370-838-191

Query Match
Best Local Similarity 4.9%; Score 151; DB 4; Length 704;
Matches 111; Conservative 71; Mismatches 190; Indels 150; Gaps 20;

QY 56 QFEAVLQHLKRSRGLATAAAIKQAAGFASKTETEPVFWYVYKVLNKHLELQRFYSLRH 115
Db 143 QFFVMEHCLK--HGLKVKKSFIGNKSFGLVLEKLCPEASDIATS-----VRNLPE 195
QY 116 IASDVGRGAWLRCAINLHSLERYLHMLLADRCRLSTFYEDWSFVMDERSMPLPTMAAG 175
Db 196 LKTAVGRGAWLYLALMOKKLADYLKVLIDNKHLLSEFYEPALMMEEB--GMVIVGLLVG 254
QY 176 LNSILFAINIDNKDLNGQ-----SKFAPTVDL--LKESTQNTVSLKKESTQGVSSLFR 227
Db 255 LNVLDANLCKGEDLDSQGVDFSLYKDVQDLGGKEH-ERITDVLQKNY-----306
QY 228 EITASSAVSILIKPEQTDPCSCPGMSVLMFNNAKSGRRKXKPTXSHLMRKMRTILG 287
Db 307 -----VEELNRHLSC---TVGDLQTKIDGLEK---TNSKLOEELSAAADR 345
QY 288 TCLKRLHGGQRAQTTPAPLISCPPLKAPSGLTPMSEEQQLMENXFPVFERGVWVPE- 346
Db 346 IC-----SLOEQOQLREQNELIRERSEKSEVEI 373
QY 347 -----AXCEKHXXRCG-XKRRXRVWKLIRKQAQPLGVAREATGREHLPLDQALG-- 397
Db 374 TKQDTKVELETYKTRQGLDEMYSVMKQKBEKKVRLLEKEL-----ELQIGMK 424
QY 398 -SAGAAQXLRHPLPCWQRLGLOPSRCPKPGERDTRGPRSPGWSWTSVQCGSLSRPR 456
Db 425 TEMEIAMKLL-----EKD-THEKQD-----TLVALRQQLVEVK 456
QY 457 KSSQPVTSASVPESM-----TISELROATVAMNRKDELEENRSLRNLDDGEMHSA 510
Db 457 AINLQMFHKAQNAESSLQKNEAITSFEGKTNQVMSMKQMEERLQHSERARQGAERSH 516
QY 511 ALRQEV-----DTLKRKV-AEQEERQGMK 533
Db 517 KIQQELGRIGALQQLSQLHEQCSLEKELKSEKQEQALQ 558

RESULT 4
US-07-741-940-7
; Sequence 7, Application US/07741940
; Patent No. 5352775
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; APPLICANT: ANAND, RAKESH
; APPLICANT: CARLSON, MARY
; APPLICANT: GRODEN, JOANNA
; APPLICANT: HEDGE, PHILIP J.
; APPLICANT: JOSLYN, GEOFF
; APPLICANT: KINZLER, KENNETH
; APPLICANT: MARKHAM, ALEXANDER F.
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: THLIVERIS, ANDREW
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4598
; COMPUTER READABLE FORM:

```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/741,940
; FILING DATE: 19920109
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.035574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2842 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: APC
; US-07-741-940-7

Query Match          4.4%; Score 136.5; DB 1; Length 2842;
Best Local Similarity 21.9%; Pred. No. 0.0014;
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;

QY 156 DWSFVMDERSSMLPTMAAGLNSILFAINDKDLNGQSKFAPTVSDLLKESTONVTSLL 215
DB 2099 DWKAIQE-----GANSIVSLH-----QAAAAACLRSQASSDSLSL- 2137
QY 216 KESTQGVSSLFREITASSAVSILIKPEQETDPCLSGFCGMSVLMNNAK----- 262
DB 2138 -KSGISILGSPFH-----LTPDQEEKPFTSNKGRPRILKPGKSTLETCKIESSEK 2185
QY 263 --RSGR-----KKKPTXSHLMRMKMSRTLTGCLKHLGQGRAQT 302
DB 2186 GIKGGKKVYKSLITGKVRNSSEISGQMKQPLQANMPSISRGTN-----IHI-PGVNNS 2239
QY 303 TPTAPLSICPPLKAPSGLTPMESEQQLMENXFPVFERGVWVPEACEKHXRCGXKRR 362
DB 2240 SSTSPVSKGPPPLKTPASKSPSEQQTATTS-----PRGA-KPSVKSE----- 2280
QY 363 XRVWKLIRKBAQGLGVAREATGREHLPLDPAQLGSAEG-----AAQXLRHPLPCQWRGL 417
DB 2281 --LSPVARQTSQ--IGGSSKAPRS-----GSRDSTPSRPAQQLSRPIQSPGRNS 2327
QY 418 LQPSR---CPPRKPGEDRTRGPRSPGWSVQCG--SOLSRPKSSEQPV----- 464
DB 2328 ISPCRNISPPNKLSQLPRTSSP-STASTKSSGSKMSYTPSGQMSQQNLTKQTGLSKN 2386
QY 465 SASVPESMITSE-LRQATVAM-MNRKDEL-----EENRSLNLLDGENHSAAL 512
DB 2387 ASSIPRESASKGLNQNNNGNKGANKVELSRMSSTKSSGESDRSERPVL---VRQSTFI 2443
QY 513 RQEVDT-TLKRKVAE 525
DB 2444 KEAPSPTLRRKLEE 2457

RESULT 5
US-08-289-548A-7
; Sequence 7, Application US/08289548A
; Patent No. 5648212
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; APPLICANT: ANAND, RAKESH
; APPLICANT: CARLSON, MARY
```

```
; APPLICANT: GRODEN, JOANNA
; APPLICANT: HEDGE, PHILIP J.
; APPLICANT: JOSLYN, GEOFF
; APPLICANT: KINZLER, KENNETH
; APPLICANT: MARKHAM, ALEXANDER F.
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: THLIVERIS, ANDREW
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, LTD
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4598
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/289,548A
; FILING DATE: 12-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.46943
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2842 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: APC
; US-08-289-548A-7

Query Match          4.4%; Score 136.5; DB 1; Length 2842;
Best Local Similarity 21.9%; Pred. No. 0.0014;
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;

QY 156 DWSFVMDERSSMLPTMAAGLNSILFAINDKDLNGQSKFAPTVSDLLKESTONVTSLL 215
DB 2099 DWKAIQE-----GANSIVSLH-----QAAAAACLRSQASSDSLSL- 2137
QY 216 KESTQGVSSLFREITASSAVSILIKPEQETDPCLSGFCGMSVLMNNAK----- 262
DB 2138 -KSGISILGSPFH-----LTPDQEEKPFTSNKGRPRILKPGKSTLETCKIESSEK 2185
QY 263 --RSGR-----KKKPTXSHLMRMKMSRTLTGCLKHLGQGRAQT 302
DB 2186 GIKGGKKVYKSLITGKVRNSSEISGQMKQPLQANMPSISRGTN-----IHI-PGVNNS 2239
QY 303 TPTAPLSICPPLKAPSGLTPMESEQQLMENXFPVFERGVWVPEACEKHXRCGXKRR 362
DB 2240 SSTSPVSKGPPPLKTPASKSPSEQQTATTS-----PRGA-KPSVKSE----- 2280
QY 363 XRVWKLIRKBAQGLGVAREATGREHLPLDPAQLGSAEG-----AAQXLRHPLPCQWRGL 417
DB 2281 --LSPVARQTSQ--IGGSSKAPRS-----GSRDSTPSRPAQQLSRPIQSPGRNS 2327
QY 418 LQPSR---CPPRKPGEDRTRGPRSPGWSVQCG--SOLSRPKSSEQPV----- 464
DB 2328 ISPCRNISPPNKLSQLPRTSSP-STASTKSSGSKMSYTPSGQMSQQNLTKQTGLSKN 2386
```

QY	465	SASVPESTMWISL-LRQATVAM-MNRKDEL	-----EENRSLRNLIDGEMHSAAL	512
Db	2387	ASSIPRSESASGLNQMNGANGANKVELSRMSSTKSSGSDRSRFPVL	-----VRQSTFI	2443
QY	513	ROEVD-TLKRKVAE	525	
Db	2444	KEAPSTLRRKLEE	2457	
RESULT 6				
US-08-452-654-7				
; Sequence 7, Application US/08452654				
; Patent No. 5691454				
; GENERAL INFORMATION:				
; APPLICANT: ALBERTSEN, HANS				
; APPLICANT: ANAND, RAKESH				
; APPLICANT: CARLSON, MARY				
; APPLICANT: GRODEN, JOANNA				
; APPLICANT: HEDGE, PHILIP J.				
; APPLICANT: JOSLYN, GEOFF				
; APPLICANT: KINZLER, KENNETH				
; APPLICANT: MARKHAM, ALEXANDER F.				
; APPLICANT: NAKAMURA, YUSUKE				
; APPLICANT: THLIVERIS, ANDREW				
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC				
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS				
; NUMBER OF SEQUENCES: 94				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: Banner, Birch, McKie & Beckett				
; STREET: 1001 G Street, NW				
; CITY: Washington				
; STATE: D.C.				
; COUNTRY: USA				
; ZIP: 20001-4598				
; COMPUTER READABLE FORM:				
; COMPUTER: IBM PC compatible				
; OPERATING SYSTEM: PC-DOS/MS-DOS				
; SOFTWARE: Patent In Release #1.0, Version #1.25				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/08/452,654				
; FILING DATE: 25-MAY-1995				
; CLASSIFICATION: 536				
; PRIOR APPLICATION DATA:				
; APPLICATION NUMBER: US 07/741,940				
; FILING DATE: 08-AUG-1991				
; ATTORNEY/AGENT INFORMATION:				
; NAME: Kagan, Sarah A.				
; REGISTRATION NUMBER: 32,141				
; REFERENCE/DOCKET NUMBER: 1107.035574				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: 202-508-9100				
; TELEFAX: 202-508-9299				
; INFORMATION FOR SEQ ID NO: 7:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 2842 amino acids				
; TYPE: amino acid				
; STRANDEDNESS: single				
; TOPOLOGY: linear				
; MOLECULE TYPE: protein				
; ORIGINAL SOURCE:				
; ORGANISM: Homo sapiens				
; IMMEDIATE SOURCE:				
; CLONE: APC				
US-08-452-654-7				
Query Match 4.4%; Score 136.5; DB 1; Length 2842;				
Best Local Similarity 21.9%; Pred. No. 0.0014;				
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;				
QY	156	DWSFVMDERSSMLPTWAAGLSILFAINTDKDLNGQSKFAFTVSDLLKESQNTVTSLL	215	

APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSLYN, GEOFF
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THLIVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, LTD
STREET: 1001 G Street, NW
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/289,548A
FILING DATE: 12-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107.46943
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2843 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-289-548A-2

Query Match 4.4%; Score 136.5; DB 1; Length 2843;
Best Local Similarity 21.9%; Pred. No. 0.0014;
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;
QY 156 DWSFVMDERSMLPTMAAGLSILFAINDKDLNGQSKFAPTVSDLLKSTQNVTSIL 215
DB 2100 DWKAIQ-----GANSIVSSLH-----QAAAAACLSRQASSDSDSILSL- 2138
QY 216 KESTQGVSLFREITASSAVSILIKPEQETDPCLSGPGMSVLMPNAK----- 262
DB 2139 -KSGISLGSPPH-----LTPDQEKFPSTNKGPRILKPGKESTLETTKIESSEK 2186
QY 263 --RSGR-----KRKXPTXSHLMRMKMSRTLTCLKRLHIGQGRAQT 302
DB 2187 GIKGGKVKYSLITGKVRNSRISGQMKQPLQANMPSISRGRTM-----IHI-PGVRNSS 2240
QY 303 TTPATLISCPKAPSGLTPTNESQOLMENXFPVPERGVWPEACEKHXRCGXRR 362
DB 2241 SSTSVPKKGPPKLPKTPASKSPSEGGTATTS-----PRGA-KPSVKSE----- 2281
QY 363 XRVKWLIRKAGQGLGVAREATGREHLPLPDAQLGSAEG-----AAQXLRHPLPCQWRGL 417
DB 2282 --LSPVARTSQ--TGGSKAPRS-----GSRDSTPSPAPQPLSRPIQSPGRNS 2328
QY 418 LQPSR---CPPKPGEDTRGPRSPGWSVTSVQCG--SOLSRPKSSBPVFT----- 464
DB 2329 ISFGNGISPPNKLSQLPRTSSP-STASTKSSGSGKMSYTSPPGRQMSQONLTKQTGLSKN 2387
QY 465 SASVPESMTISE-LQATVAM-MNRKDEL-----EENRSIRLLDGEHEHSAAL 512
DB 2388 ASSIPRSESASKGLNQMNNGNANKKVELSRMSTKSGSGESDRSERPVL-----VRQSTFI 2444

QY 513 RQEVD-TLKRKVAE 525
DB 2445 KEAPSTLRRKLEE 2458
RESULT 10
US-08-452-654-2
Sequence 2, Application US/08452654
Patent No. 5691454
GENERAL INFORMATION:
APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSLYN, GEOFF
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THLIVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, NW
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,654
FILING DATE: 25-MAY-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/741,940
FILING DATE: 08-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107.035574
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2843 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-452-654-2

Query Match 4.4%; Score 136.5; DB 1; Length 2843;
Best Local Similarity 21.9%; Pred. No. 0.0014;
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;
QY 156 DWSFVMDERSMLPTMAAGLSILFAINDKDLNGQSKFAPTVSDLLKSTQNVTSIL 215
DB 2100 DWKAIQ-----GANSIVSSLH-----QAAAAACLSRQASSDSDSILSL- 2138
QY 216 KESTQGVSLFREITASSAVSILIKPEQETDPCLSGPGMSVLMPNAK----- 262
DB 2139 -KSGISLGSPPH-----LTPDQEKFPSTNKGPRILKPGKESTLETTKIESSEK 2186
QY 263 --RSGR-----KRKXPTXSHLMRMKMSRTLTCLKRLHIGQGRAQT 302
DB 2187 GIKGGKVKYSLITGKVRNSRISGQMKQPLQANMPSISRGRTM-----IHI-PGVRNSS 2240


```

;
; TITLE OF INVENTION: MUTATIONS IN THE APC GENE
;
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER & WITCOFF, LTD.
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20001
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,235A
; FILING DATE: 01-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.48688
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 508 9100
; TELEFAX: 202 508 9299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2843 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-370-235A-2
;
; Query Match 4.4%; Score 136.5; DB 2; Length 2843;
; Best Local Similarity 21.9%; Pred. No. 0.0014;
; Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;
;
; QY 156 DWSFVMDERSMLPTMAAGLSILPAINDKDLNGQSKFAPTVDLLKESTQNTVLSL 215
; DB 2100 DWKAIQE-----GANSIVSSLH-----QAAAAACLSRQASDSDSILSL- 2138
;
; QY 216 KESTQGVSSLFREITASSAVSILIKPEQETDPCLSCPGMSVLMENAK----- 262
; DB 2139 -KSGISLGSPFH-----LTPDQEEKPFTSNKGPRILKPGKEKSTLETKKIESK 2186
;
; QY 263 --RSGR-----KRKXPTYSHLMRMKMSRTLTGCLKHLGQGRQRT 302
; DB 2187 GIKGKKVYKSLITGVRNSEISGQMKQPLQANMPSISRGRTM-----IHI-PGVRNS 2240
;
; QY 303 TPTAPLSICPPLKAPSLGTPMSEEQQLMENKFPFVFERGVWVPEAXCEKHXRCGXKR 362
; DB 2241 SSTSPVSKGPPKLTTPASKSPSEGQTATTS-----PRGA-KPSVKE----- 2281
;
; QY 363 XRVWKLIRKEAQGLGVAREATGREHLPLPDAQLGSAG-----AAQXLRHPLPCQWRGL 417
; DB 2282 --LSPVARQTSQ--IGSSKAPSKS-----GSRDSTPSRPAQQLSRPTQSPGRNS 2328
;
; QY 418 LQPSR---CPPRKGERDTRGPRSPGWSVQCG--SOLSRPKSSEQPV----- 464
; DB 2329 ISPGRNGISPPNKLSQLPRTSSP-STASTKSSGGKMSYTPGROMSQOQLTKQTGLSKN 2387
;
; QY 465 SASVPESMTISE-LRQATVAM-MNRKDEL-----EENRSLRNLDDGMEHSAAL 512
; DB 2388 ASSIPRESASKGLNQMNNGANGANKVELSRMSSTKSSGSESDRSERPVL---VRQSTFI 2444
;
;
; RESULT 14
; US-08-450-582-2
; Sequence 2, Application US/08450582
;
;
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/289,548
; FILING DATE: 12-AUG-1994
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US 07/741,940
; FILING DATE: 08-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.49964
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2843 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; US-08-452-655B-7
;
; Query Match 4.4%; Score 136.5; DB 1; Length 2843;
; Best Local Similarity 21.9%; Pred. No. 0.0014;
; Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;
;
; QY 156 DWSFVMDERSMLPTMAAGLSILPAINDKDLNGQSKFAPTVDLLKESTQNTVLSL 215
; DB 2100 DWKAIQE-----GANSIVSSLH-----QAAAAACLSRQASDSDSILSL- 2138
;
; QY 216 KESTQGVSSLFREITASSAVSILIKPEQETDPCLSCPGMSVLMENAK----- 262
; DB 2139 -KSGISLGSPFH-----LTPDQEEKPFTSNKGPRILKPGKEKSTLETKKIESK 2186
;
; QY 263 --RSGR-----KRKXPTYSHLMRMKMSRTLTGCLKHLGQGRQRT 302
; DB 2187 GIKGKKVYKSLITGVRNSEISGQMKQPLQANMPSISRGRTM-----IHI-PGVRNS 2240
;
; QY 303 TPTAPLSICPPLKAPSLGTPMSEEQQLMENKFPFVFERGVWVPEAXCEKHXRCGXKR 362
; DB 2241 SSTSPVSKGPPKLTTPASKSPSEGQTATTS-----PRGA-KPSVKE----- 2281
;
; QY 363 XRVWKLIRKEAQGLGVAREATGREHLPLPDAQLGSAG-----AAQXLRHPLPCQWRGL 417
; DB 2282 --LSPVARQTSQ--IGSSKAPSKS-----GSRDSTPSRPAQQLSRPTQSPGRNS 2328
;
; QY 418 LQPSR---CPPRKGERDTRGPRSPGWSVQCG--SOLSRPKSSEQPV----- 464
; DB 2329 ISPGRNGISPPNKLSQLPRTSSP-STASTKSSGGKMSYTPGROMSQOQLTKQTGLSKN 2387
;
; QY 465 SASVPESMTISE-LRQATVAM-MNRKDEL-----EENRSLRNLDDGMEHSAAL 512
; DB 2388 ASSIPRESASKGLNQMNNGANGANKVELSRMSSTKSSGSESDRSERPVL---VRQSTFI 2444
;
;
; RESULT 13
; US-08-370-235A-2
; Sequence 2, Application US/08370235A
; Patent No. 5910418
; GENERAL INFORMATION:
; APPLICANT: VOGELSTEIN, BERT
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: HILL, DAVID E.
; APPLICANT: JOHNSON, KAREN A.
; TITLE OF INVENTION: ANTIBODIES AND ASSAYS FOR DETERMINING
```


Mon Jun 13 13:27:10 2005

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
US-08-450-582-7

Query Match          4.4%; Score 136.5; DB 3; Length 2843;
Best Local Similarity 21.9%; Pred. NO. 0.0014;
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;

QY 156 DWSFVWDEERSMLFTMAAGLNSILFAINIDKDLNGQSKFAPTVDLLKKESTQNTSLL 215
Db 2100 DWKAIQE-----GANSIVSSLH-----QAAAAACLRSQASSDSLSL- 2138
QY 216 KESTQGVSLFREITASSAVSILIKPEQETDPCLSQPCQMSVLMPNAK----- 262
Db 2139 -KSGISLGSPPH-----LTPDQEKPFSTNKGPRILKPGCEKSTLETKKIESEK 2186
QY 263 --RSGR-----KRXPTYXSHLMRKMSRTLTGCLKRHLGQQAQRT 302
Db 2187 GIKGKKVYKSLITGKVRNSNISEIGMKOPLQANMPSISGRITM-----IHI-PGVRNSS 2240
QY 303 TPTAPLISCPPLKAPSGLTPMSEQQLMENXFPVFERGVVWVPEAXCEKHXRCGXKRR 362
Db 2241 SSTSPVKKGPPPLKTPASKSPSEGTATTS-----PRGA-KPSVKSE----- 2281
QY 363 XRVWKLIRKEAQOGLGVAREATGREHLPLDQQLGSAEG-----AAQXLRHPLPCQWRGL 417
Db 2282 --LSPVARQTSQ--IGGSKAPSR-----GSRDSTPSRPAQQLSRFIQSPGRNS 2328
QY 418 LQPSR---CPPRXGERDTRGPRSPGWSVQCG--SOLSRPKSSQOPVT----- 464
Db 2329 ISFGRNGISPPNKLSQLPTSSP-STASTKSGSGGWMYTSFGRQMSQNLTKQTGLSKN 2387
QY 465 SASVPESMTISE-LRQATVAM-MNRKDEL-----EENRSLRNLLDGENEHSAL 512
Db 2388 ASSIPRESASAKGLNQNNNGANGKKVELSRMSSTKSGSGESDRSRPVL---VRQSTFI 2444
QY 513 RQEVDTLKRKVAE 525
Db 2445 KEAPSTLRRKLEE 2458
```

Search completed: June 7, 2005, 12:04:48
Job time : 25.6321 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 12:02:31 ; Search time 76.4207 Seconds
(without alignments)
2846.907 Million cell updates/sec

Title: US-09-155-676B-2
Perfect score: 3093
Sequence: 1 XTGPXGXMGSMXNDXKRF.....FLWTFHVCEPINCFSLSLKKK 604

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599520 seqs, 360203123 residues

Total number of hits satisfying chosen parameters: 1599520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*

2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*

3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*

5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*

6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*

7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*

9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*

10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*

11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*

12: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*

13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

17: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

18: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

19: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

20: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

21: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	183	5.9	1019	15	US-10-276-774-2298
2	175.5	5.7	473	15	US-10-094-749-1936
3	175.5	5.7	473	16	US-10-408-765A-2491
4	160	5.2	2845	15	US-10-267-502-372
5	151	4.9	600	15	US-10-015-115-73
6	151	4.9	704	9	US-09-738-973-191
7	151	4.9	704	9	US-09-854-133-191
8	151	4.9	704	14	US-10-144-649A-191
9	145	4.7	546	15	US-10-264-237-2780
10	137	4.4	286	15	US-10-276-774-1412
11	137	4.4	286	15	US-10-296-115-1340
12	137	4.4	286	15	US-10-296-115-1465

13	136.5	4.4	369	15	US-10-264-237-2045	Sequence 2045, Ap
14	136.5	4.4	2843	8	US-08-681-219-32	Sequence 32, Appl
15	136.5	4.4	2843	9	US-09-987-482-1	Sequence 1, Appl
16	136.5	4.4	2843	10	US-09-230-111C-30	Sequence 30, Appl
17	136.5	4.4	2843	14	US-10-092-138-30	Sequence 30, Appl
18	136.5	4.4	2843	15	US-10-392-113-21	Sequence 21, Appl
19	136.5	4.4	2843	16	US-10-408-765A-1970	Sequence 1970, Ap
20	136.5	4.4	2843	16	US-10-820-403-30	Sequence 370, App
21	136.5	4.4	2844	15	US-10-267-502-370	Sequence 14, Appl
22	135.5	4.4	468	14	US-10-147-026-14	Sequence 4, Appl
23	124	4.0	494	14	US-10-035-343-4	Sequence 3547, Ap
24	124	4.0	745	15	US-10-108-260A-3547	Sequence 3340, Ap
25	124	4.0	787	15	US-10-104-047-3340	Sequence 2, Appl
26	124	4.0	1253	15	US-10-334-143-2	Sequence 126375,
27	123.5	4.0	1274	16	US-10-437-963-126375	Sequence 81, Appl
28	123	4.0	827	15	US-10-428-275-234	Sequence 234, App
29	123	3.9	827	15	US-10-428-275-234	Sequence 810, App
30	121.5	3.9	790	15	US-10-225-066A-810	Sequence 2384, Ap
31	121.5	3.9	790	15	US-10-374-780A-2384	Sequence 33, Appl
32	121	3.9	666	15	US-10-168-067C-33	Sequence 1, Appl
33	121	3.9	1190	15	US-10-168-067C-1	Sequence 61, Appl
34	120.5	3.9	1608	11	US-09-964-956-61	Sequence 60, Appl
35	120.5	3.9	1788	11	US-09-964-956-60	Sequence 63, Appl
36	120.5	3.9	2641	11	US-09-964-956-63	Sequence 1425, Ap
37	120.5	3.9	2641	16	US-10-408-765A-1425	Sequence 27, Appl
38	120.5	3.9	2811	11	US-09-964-956-27	Sequence 25, Appl
39	120.5	3.9	2814	11	US-09-964-956-25	Sequence 45308, A
40	119.5	3.9	520	15	US-10-425-114-45308	Sequence 183619,
41	118	3.8	862	16	US-10-437-963-183619	Sequence 5585, Ap
42	118	3.8	1246	15	US-10-369-493-6585	Sequence 226, App
43	117	3.8	303	15	US-10-428-275-236	Sequence 230, App
44	117	3.8	827	15	US-10-428-275-222	
45	117	3.8	827	15	US-10-428-275-230	

ALIGNMENTS

RESULT 1

US-10-276-774-2298
; Sequence 2298, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2298
; LENGTH: 1019
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-2298

Query Match 5.9%; Score 183; DB 15; Length 1019;
Best Local Similarity 22.2%; Pred. No. 1.3e+06;
Matches 115; Conservative 75; Mismatches 196; Indels 132; Gaps 25;

Qy	21	LLERLLDAVKQCIQIRFXG-RKEITAS--DSDRVTCLCAQFENVLQGLKESRGLATAAA	77
Db	10	ILENISLVKLGQSYFAACEDEIPAIRNHDKVLQRLCELDHALLYLGLQ-----	58
Qy	78	IKQAAGFASKTETPEPVFVYVVKLVNKHLELQRFVSLRHIAADVGRGRAWLRCALEHLSLE	137
Db	59	-----DLSSGIWLVLVHFTREAIKQIEVLQVATNLRGSRRAWLYLAUNENSLE	107


```

RESULT 10
US-10-276-774-1412
; Sequence 1412, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq. Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: NO. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700

```


Db 2282 --LSPVARQTSQ--ICGSSKAPRS-----GSRDSTPSRPAQQPLSRPIQSPGRNS 2328
QY 418 LQPSF---CPPKPGERDRTRGPRSPGSGWTSVQCG--SOLSRPRKSSQPVT----- 464
Db 2329 ISPGRNGISPPNKLSQLPRTSSP-STASTKSGSGKMSYTPGROMSQOQNLTKQTGLSKN 2387
QY 465 SASVPESMTISE-LRQATVAM-MNRKDEL-----EENRSLRNLDDGEMEHSAAAL 512
Db 2388 ASSIERSESASKGLNQMNNGCANKKVELSRMSSTKSSGSESDRSERPVL---VRQSTFI 2444
QY 513 ROEVD-TLKRKVAE 525
Db 2445 KEAPSTLRRKLEE 2458

Search completed: June 7, 2005, 12:25:31
Job time : 78.4207 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 22:58:14 ; Search time 894.226 Seconds
(without alignments)
17417.120 Million cell updates/sec

Title: US-09-155-676B-3
Perfect score: 2631
Sequence: 1 ccctctcacagccagcc.....acgtcgacctcgaggggg 2631

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2553.8	97.1	2631	2	AAV03324
2	1851.2	70.4	4596	2	AAV03326
3	1850.2	70.3	4596	5	AA579490
4	1850.2	70.3	4596	6	ABK51171
5	1850.2	70.3	4596	10	AD85075
6	1850.2	70.3	4596	13	ADR25114
7	1783.4	67.8	4684	11	ACN44439
8	1771.6	67.3	3152	12	ADQ67410
9	1133.4	43.1	3156	2	AAV73917
10	1133.4	43.1	3156	2	AAV71603
11	1133.4	43.1	3156	2	AAV69285
12	977	37.1	2844	2	AA587842
13	973.4	37.0	73967	11	ACN44438
14	754.2	28.7	1385	4	AAH99262
15	691.8	26.3	2760	5	AA579491
16	672.8	25.6	2828	11	ACN44437
17	588.2	22.4	722	5	AA579492
18	446.8	17.0	2070	12	ADP28582
19	349.4	13.3	2475	11	ADM03069
20	242.2	9.2	499	9	ACH15373

21	205.8	7.8	476	10	ADF79802
22	188.8	7.2	42566	11	ACN44436
C 23	88.4	3.4	455	5	ABV22061
C 24	88.4	3.4	455	5	ABV27896
C 25	88.4	3.4	476	5	ABV12334
C 26	88.4	3.4	504	5	ABV33478
C 27	88.4	3.4	504	5	ABV42401
C 28	73.8	2.8	463	5	ABV03165
C 29	70	2.7	1024	5	AA579493
C 30	58.2	2.2	2000	8	ADA71938
C 31	53.2	2.0	1253	2	AAV03325
C 32	52.8	2.0	65	6	ABN52988
C 33	50.4	1.9	374	4	AAI84793
C 34	50	1.9	50	6	ABZ03850
C 35	50	1.9	50	6	ABZ02592
C 36	49	1.9	60	6	ABN46309
C 37	43.6	1.7	2000	8	ADA71938
C 38	42.6	1.6	1464	6	AD24766
C 39	42.4	1.6	825	11	ABD10419
C 40	42.4	1.6	1359	11	ABD10504
C 41	42.4	1.6	1746	11	ABD10381
C 42	42	1.6	5280	12	ADN89158
C 43	41.2	1.5	1097	10	ADJ92320
C 44	40.4	1.5	1089	8	ACA26986
C 45	40.4	1.5	1711	1	AA60310

ALIGNMENTS

RESULT 1
AAV03324
ID AAV03324 standard; cDNA; 2631 BP.
XX
AC AAV03324;

DT 15-APR-1998 (first entry)

DE Clone 10 cDNA encoding NMP1, a TRAF2 binding protein.

XX Human tumour necrosis factor receptor-associated factor 2; TRAF2;
KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;
KW intracellular signalling activity; acute hepatitis;
KW autoimmune-induced cell death; ss.

OS Homo sapiens.

XX WO9737016-A1.

XX 09-OCT-1997.

XX 01-APR-1997; 97WO-IL000117.

XX 02-APR-1996; 96IL-00117800.

XX 26-AUG-1996; 96IL-00119133.

XX (YEDA) YEDA RES & DEV CO LTD.

PI Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;

XX WPI; 1997-503101/46.

PT DNA encoding tumour necrosis factor receptor-associated factor binding
PT molecule - used for modulation or mediation in cells of the activity of
PT NF-KB.

XX Claim 4; Fig 4; 127pp; English.

XX The present sequence encodes a protein designated NMP1, a TRAF2 binding
CC protein. This protein contains Ser/Thr protein kinase motifs. A full
CC length cDNA clone encoding NF-kappaB (NIK) was obtained by PCR using the
CC present clone, clone 10. The clone 10 protein is capable of binding to at
CC least amino acids 222-501 of TRAF2. The TRAF-2 binding proteins can be

CC used for modulation or mediation in cells of the activity of NF-kappaB or
 CC any other intracellular signalling activity modulated or mediated by
 CC TRAF2. TRAF-binding proteins are especially used for prevention or
 CC treatment of pathological conditions associated with NF-kB induction,
 CC e.g. acute hepatitis, autoimmune-induced cell death, e.g. death of the
 CC beta langerhans cells or the pancreas that results in diabetes, the death
 CC of cells in graft rejection, the death of oligodendrocytes in the brain
 CC in multiple sclerosis, and AIDS-inhibited T cell suicide which causes
 CC proliferation of the AIDS virus and hence the AIDS disease. The proteins
 CC are also useful for screening of ligands capable of binding to a protein,
 CC which are useful for modulating cellular activity modulated/mediated by
 CC TRAF2

XX	Sequence	2631 BP;	596 A;	791 C;	704 G;	459 T;	0 U;	81 Other;	
QY	Query Match	97.1%;	Score 2553.8;	DB 2;	Length 2631;				
QY	Best Local Similarity	100.0%;	Pred. No. 0;						
QY	Matches 2631;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	CCCCCTCTCAGCCAGCCAGCCATCCAGAGGGCTGAGGAAGAGCCCATCCACCGCGTGT	60						
QY	1	CCCCCTCTCAGCCAGCCAGCCATCCAGAGGGCTGAGGAAGAGCCCATCCACCGCGTGT	60						
QY	61	CTGCGAGCGAGCTGGGAGGAGGAGTGAACCGGGCCTACAGCAAGTGGGAGGTCTGAAGA	120						
QY	61	CTGCGAGCGAGCTGGGAGGAGGAGTGAACCGGGCCTACAGCAAGTGGGAGGTCTGAAGA	120						
QY	121	GGCCTTGGAGGGGAGGAATATAAGAACCAAGACATCCACCGCAATTAAC	180						
QY	121	GGCCTTGGAGGGGAGGAATATAAGAACCAAGACATCCACCGCAATTAAC	180						
QY	181	ACCAGACCTCTCAGCCAGCCAGAGAGCTTTCGCCAGGGCCCGAGGGCCCGCCAG	240						
QY	181	ACCAGACCTCTCAGCCAGCCAGAGAGCTTTCGCCAGGGCCCGAGGGCCCGCCAG	240						
QY	241	CTGAGGAGACAA CAGGCGAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAG	300						
QY	241	CTGAGGAGACAA CAGGCGAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAG	300						
QY	301	AGCCAAACAGTCTCTCCCTTGACTTTGAGCAGAGGAGTCTGGGATGTGGGAACCT	360						
QY	301	AGCCAAACAGTCTCTCCCTTGACTTTGAGCAGAGGAGTCTGGGATGTGGGAACCT	360						
QY	361	TACCTCTGCTCTCCCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAG	420						
QY	361	TACCTCTGCTCTCTCCCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAG	420						
QY	421	CAACCGTCCCGGAGCAGGAATCGAGAGCTGGAAATAGAAATATTCCTCAACAGGCTGT	480						
QY	421	CAACCGTCCCGGAGCAGGAATCGAGAGCTGGAAATAGAAATATTCCTCAACAGGCTGT	480						
QY	481	CCAGCCATTTCTCTGGAGGAGCAGGACAAATTCCTGCTCCCTCAGCATCGAGCC	540						
QY	481	CCAGCCATTTCTCTGGAGGAGCAGGACAAATTCCTGCTCCCTCAGCATCGAGCC	540						
QY	541	TCTCCCTGTGGATGACAGTGAGAGAAACCCATCAAGGCTCTCAAGCTCCCGGGA CA	600						
QY	541	TCTCCCTGTGGATGACAGTGAGAGAAACCCATCAAGGCTCTCAAGCTCCCGGGA CA	600						
QY	601	CCCTGAGCTCAGGCGTACA CTCTGGAGAGCAGGCGCGAGGCTCGAAGCTCCAGCTGGA	660						
QY	601	CCCTGAGCTCAGGCGTACA CTCTGGAGAGCAGGCGCGAGGCTCGAAGCTCCAGCTGGA	660						
QY	661	ACATGGTGTGGCCCGGGGGCCACCCAGACCCCAAGCTATTTCAATGGTGTGAAAG	720						
QY	661	ACATGGTGTGGCCCGGGGGCCACCCAGACCCCAAGCTATTTCAATGGTGTGAAAG	720						
QY	721	TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCGCGGAGTTCCACCGGGTCAAAG	780						
QY	721	TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCGCGGAGTTCCACCGGGTCAAAG	780						
QY	781	TGGGAGACATCGCCACTGGCATCAGAGCCAGATCCAGCTGCGAGCTTCAGCTTGTGTCA	840						

Db	1081	TTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCGAGCCTTCAGCTTGGTCA	840
QY	841	CCAAAGACGGGACGCTGTTGCTGCTAGCATGAGGTGCGAGCTCGGCGATCGACCTGC	900
Db	841	CCAAAGACGGGACGCTGTTGCTGCTAGCATGAGGTGCGAGCTCGGCGATCGACCTGC	900
QY	901	AGTGCACATCGGCGCTGATGAGCTTGGCTGAGAGTGGAGGTCAAGCATGGCCAGC	960
Db	901	AGTGCACATCGGCGCTGATGAGCTTGGCTGAGAGTGGAGGTCAAGCATGGCCAGC	960
QY	961	TGAGAAACAGGCGCTTACCTGCTCCACCGCGCTCCACACTGCGGAAAGACGACCT	1020
Db	961	TGAGAAACAGGCGCTTACCTGCTCCACCGCGCTCCACACTGCGGAAAGACGACCT	1020
QY	1021	TCCTGCTGCTGACGATGCTGCCCTGAAACACAGGCTCAGCGCTTCCAGGGGATYTG	1080
Db	1021	TCCTGCTGCTGACGATGCTGCCCTGAAACACAGGCTCAGCGCTTCCAGGGGATYTG	1080
QY	1081	NCCAGCCCCCGGCTCARCAGNTGGGAAACAGGCGCTCGNCAGCAGCNAAGTNGGGG	1140
Db	1081	NCCAGCCCCCGGCTCARCAGNTGGGAAACAGGCGCTCGNCAGCAGCNAAGTNGGGG	1140
QY	1141	CAAGCAGAGATGCTCCAGGATTTTACANCTTGAGCCCTGCCANCCCTGCTGAADA	1200
Db	1141	CAAGCAGAGATGCTCCAGGATTTTACANCTTGAGCCCTGCCANCCCTGCTGAADA	1200
QY	1201	AAACAYTNCCGCCACGTGAAGAGACAGAGGAGTGCNAGGAGTNNACCTTGGGGAA	1260
Db	1201	AAACAYTNCCGCCACGTGAAGAGACAGAGGAGTGCNAGGAGTNNACCTTGGGGAA	1260
QY	1261	ACAAAACAGGATCTTNTTCTGCCCCCTGCTCCAGTTCGAGTTCGACCTGACCTGG	1320
Db	1261	ACAAAACAGGATCTTNTTCTGCCCCCTGCTCCAGTTCGAGTTCGACCTGACCTGG	1320
QY	1321	ANTCAGTGACCATTTGTTGGCAGANCAAGGAGAGAGCTTCCAGCCTGGGTCAAGAGG	1380
Db	1321	ANTCAGTGACCATTTGTTGGCAGANCAAGGAGAGAGCTTCCAGCCTGGGTCAAGAGG	1380
QY	1381	GTGGCGAGCCCTTGGCCCTCAACCTNCCAGGCTGCTGAGAGTGTCAAGTGTGTA	1440
Db	1381	GTGGCGAGCCCTTGGCCCTCAACCTNCCAGGCTGCTGAGAGTGTCAAGTGTGTA	1440
QY	1441	AGGNCNCCAAANCTCAGGNTTCAGTGCAAGAACAGGAGTTCAGCCTGGGTCAAGAGG	1500
Db	1441	AGGNCNCCAAANCTCAGGNTTCAGTGCAAGAACAGGAGTTCAGCCTGGGTCAAGAGG	1500
QY	1501	GTTTAAANNNGGGGGCCCTCTNAAACCCCTTGCCTNNGGCTNCACTNNGGCCAGCTC	1560
Db	1501	GTTTAAANNNGGGGGCCCTCTNAAACCCCTTGCCTNNGGCTNCACTNNGGCCAGCTC	1560
QY	1561	CTTTTGGGTGTAGGGGAAAGAAATGCTGAGCCCTGGGAGGCTTCCCTGGTAGAATACAC	1620
Db	1561	CTTTTGGGTGTAGGGGAAAGAAATGCTGAGCCCTGGGAGGCTTCCCTGGTAGAATACAC	1620
QY	1621	CACACTTTTTCAGTGTGTTGCAACAGAGTCTGAGTTCAGCTTCTGGTTTTCAGCAAGG	1680
Db	1621	CACACTTTTTCAGTGTGTTGCAACAGAGTCTGAGTTCAGCTTCTGGTTTTCAGCAAGG	1680
QY	1681	AAAGAGAGTGTAAAGTGAAGTGTCTGATNCCCGCAGACATGTCGCTTGTGCTGTG	1740
Db	1681	AAAGAGAGTGTAAAGTGAAGTGTCTGATNCCCGCAGACATGTCGCTTGTGCTGTG	1740
QY	1741	GCTACCATCTTCCCGCAGCAGAGGCGCCGAGCCCTTTCAGGCGCCAGCAGCTGCCCCAG	1800
Db	1741	GCTACCATCTTCCCGCAGCAGAGGCGCCGAGCCCTTTCAGGCGCCAGCAGCTGCCCCAG	1800
QY	1801	ACTCGCTGCACTTCAGTTCCTCTCATCTGTAAAGTGAAGGTGATGAGGATATGCTGTA	1860
Db	1801	ACTCGCTGCACTTCAGTTCCTCTCATCTGTAAAGTGAAGGTGATGAGGATATGCTGTA	1860
QY	1861	CAGGAAACAGTCTGTGGATGGAATGATCAGTCTTAAAGGAAAGGAGAGAGAGAGAG	1920
Db	1861	CAGGAAACAGTCTGTGGATGGAATGATCAGTCTTAAAGGAAAGGAGAGAGAGAGAG	1920

Db 2519 CAACCGTCCGGAGAGGAACTGCAGCAGCTGGAAATAGAAATTATTCTCTCAACAGCCTGT 2578
 Qy 481 CCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGCCTCAGCATCGACAGCC 540
 Db 2579 CCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGCCTCAGCATCGACAGCC 2638
 Qy 541 TCTCCCTGTGCGATGACAGTGAGAAACCCATCAAGAGCCTCTCAAGCTCGCGGAC 600
 Db 2639 TCTCCCTGTGCGATGACAGTGAGAAACCCATCAAGAGCCTCTCAAGCTCGCGGAC 2698
 Qy 601 CCTCAGCTCAGCGGTACACTCTCTGAGCAGCGCCAGGCTCGAAGCTCCAGCTGGA 660
 Db 2699 CCTCAGCTCAGCGGTACACTCTCTGAGCAGCGCCAGGCTCGAAGCTCCAGCTGGA 2758
 Qy 661 ACATGGTGTGCTCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 720
 Db 2759 ACATGGTGTGCTCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 2818
 Qy 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCGGAGTTCCACCGGTCAAAG 780
 Db 2819 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCGGAGTTCCACCGGTCAAAG 2878
 Qy 781 TGGGAGACATCGCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTCAGCTTGGTCA 840
 Db 2879 TGGGAGACATCGCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTCAGCTTGGTCA 2938
 Qy 841 CCAAGACCGGAGCCTGTTGCTAGCAGATGGAGTGCAGACTCGGGCATCGACCTGC 900
 Db 2939 CCAAGACCGGAGCCTGTTGCTAGCAGATGGAGTGCAGACTCGGGCATCGACCTGC 2998
 Qy 901 AGTGACACATGCGCCCTGATGCGAGCTTCGCTCGAGCTGGAGGTCAAGCATGSCCAGC 960
 Db 2999 AGTGACACATGCGCCCTGATGCGAGCTTCGCTCGAGCTGGAGGTCAAGCATGSCCAGC 3058
 Qy 961 TGAGAACAGGCGCTTAACCTTGCCTCCACCGCGCTCCACATGCGCGAAAGAGCCT 1020
 Db 3059 TGAGAACAGGCGCTTAACCTTGCCTCCACCGCGCTCCACATGCGCGAAAGAGCCT 3118
 Qy 1021 TCTGCTCGGTGACAGATGCTGCCCTGAAACACAGGCTCAGCGTCCAGGGATGTG 1080
 Db 3119 TCTGCTCGGTGACAGATGCTGCCCTGAAACACAGGCTCAGCGTCCAGGGATGTG 3176
 Qy 1081 NCCAGCCCCCGCTCARGAGTGGAAACAGGCGCTCGNACAGCNAGAGTGGGG 1140
 Db 3177 GCCAGCCCCCGCTCA--CAGTGGGAAACAGGCGCTCG-----CAGCAGCAAGGTGGG 3229
 Qy 1141 CAAGCNAGAAATGCTCCAGGATTCACANCTGAGCCGCTCCAGCCAGCTGCTGAADA 1200
 Db 3230 GCAAGCAGAAATGCTCCAGGATTCACACTGAGCCCTGCCCCAC-----CCTGCTGA 3283
 Qy 1201 AAACATNCCGACGTTGAAGACAGAGGAGGATGGNACGAGTTNNAACCTYGGGGAA 1260
 Db 3284 AAACATNCCGACGTTGAAGACAGAGGAGGATGGC-----AGGAGTTACCTGGGAA 3339
 Qy 1261 AAAAAACAGGATCTTTTCTGCCCCCTGCTCAGTNCAGTGGCTGAGCCGCTGAG 1320
 Db 3340 AAAAAACAGGATCTTTTCTGCCCCCTGCTCAGT--CGAGTTGGCTGA--CCCGCTTG 3395
 Qy 1321 ANTCAGTGACATTTGTTGACAGNACAGGAGAGCAGCTTCAGCTGGTTCAGAGGG 1380
 Db 3396 GATCAGTGACATTTGTTGACAG--CAGGAGAGCAGCTTCAGCTGGTTCAGAGGG 3454
 Qy 1381 GTGGGCGAGCCCTTCGCGCCCTCACCTNCCAGCTGCTGTGNAGAGTGTCAAGTGTGA 1440
 Db 3455 GTGGGCGAGCCCTTCGCGCCCTCACCT--CCAGCTGCTGTG--AGAGTGTCAAGTGTGA 3512
 Qy 1441 AGGNCACAAANTCAGGNTTCAGTGACAGAACAGGTTNAGAGGATGATCCGCCGNTA 1500
 Db 3513 AGGNCACAACTC---AGGTTCAGTGACAGAACAGGT--CAGCAGGATATGCCCCCGGTAG 3568
 Qy 1501 GGTTAANNGGGGCCCTCTNAACCCCTTGCTNGGCTNACCTNGCCGAGCTCANCCC 1560
 Db 3569 GTTAA-----GGGGGCCCTCTNAACCCCTTGCTNGGCTCA---CCTGGCCAGCTCA-CCC 3620

Qy 1561 CTTTTCGGTCTAGGGAAAAGATGCTGACCTCGGAAAGGCTTCCTGTGTAGATAACAC 1620
 Db 3621 CTTTTCGGTCTAGGGAAAAGATGCTGACCTCGGAAAGGCT--CCTGTGTAGATAACAC 3679
 Qy 1621 CACACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCTGTTGTCAGCAAGACC 1680
 Db 3680 CACACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCTGTTGTCAGCAAGACC 3739
 Qy 1681 AAAGAAGTGTGTAAGTGAAGTGGTCTCAGTNCCTCCAGACATGTCCTCTTGTGCTGCTG 1740
 Db 3740 AAAGAAGTGTGTAAGTGAAGTGGTCTCAGT--CCCAGACATGTCCTCTTGTGCTGCTG 3798
 Qy 1741 GCTACCACTCTTCCCAGACAGCAGCCCGAGCCCTTTCAGGCCACGACACTGCCCCAG 1800
 Db 3799 GCTACCACTCTTCCCAGACAGCAGCCCGAGCCCTTTCAGGCCACGACACTGCCCCAG 3858
 Qy 1801 ACTCGCTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGTGATGAGGATATGCTGTA 1860
 Db 3859 ACTCGCTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGTGATGAGGATATGCTGTA 3918
 Qy 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAGNAAAGCAGCAGAGAGACGY 1920
 Db 3919 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAA--GGAAAGCAGCAGAGAGAGC- 3975
 Qy 1921 TCCGCGCCCCCAGNCCCACCTNATCAGTGTNCCAGCCTGTCTNGGTTNCCCAGNAGACA 1980
 Db 3976 TCCGCGCCCCCAGCCTTATCAGTGT---CCAGCCTGTCTGTTCCCCAGAGCAGC- 4031
 Qy 1981 GCTNAGCATCANCACTGACATNCACCTNGCCCTGCCCCCTNGGCCANGAGGTTACTG 2040
 Db 4032 -----TCAGCATCACACTGACACTCACCTGCCCCCTGCCCCCTG--CCAGAGGTTACTG 4082
 Qy 2041 CCGNACGCACTTTCGACNTCTGATGNACCTCAAGCACTTTCATGCTNGCCCTCTTNG 2100
 Db 4083 CCG--ACGCACTTTCGACT--CTGATGACCTCAAGCACTTTCATGCTGCCCCCTCTGCA 4139
 Qy 2101 GCAGGNCAGGNCAGGNCAGTGCATCTGTAGNAGCATANGCAAGCCAGAGATGG 2160
 Db 4140 GGGCAGGCGCAGGCGAGTGACACT-----GTAGGAGCATAGCAAGCCAGAGATGG 4189
 Qy 2161 GGTGNAAGGANCACAGTCTTGAGCTGCCANCATGATGTGACTNCTCAAACTCTTN 2220
 Db 4190 GGTGAA--GGGACACAGTCTTGAGCTGTCCA--CATGATGTGACT--CCTCAACTCTT- 4244
 Qy 2221 NCCAGNATTTCTTAAGAAATAGCANCCTTNCCTTCCCAATGCCCCAGCTTACCTTCT 2280
 Db 4245 --CCAGATTTCTTAAGAAATAGCACCCTT-----TCCCAATGCCCCAGCTTACCTTCT 4300
 Qy 2281 CCAGGCGAGCTNCTCAGGACTCAGTGTAGCATTAATCAGCTGTGNAATCGTCAGGGGG 2340
 Db 4301 CCAGGCGAGCTA--CTCAGGACTCAGTGTAGCATTAATCAGCTGTG--AATCGTCAGGGGG 4358
 Qy 2341 TGTCTGCTAGCTCAACCTCTGCGGCGAGGGAGCGCCGAGACTCCGTGGGAGAGCTCAT 2400
 Db 4359 TGTCTGCTAGCTCAACCTCTGCGGCGAGGGAGCGCCGAGACTCCGTGGGAGAGCTCAT 4418
 Qy 2401 TCCCACTCTTGGCCAAAGCAGCCTTTNGTTCAGCTGTCCATTCAGTTCAGACTGCTCCC 2460
 Db 4419 TCCCACTCTTGGCCAAAGCAGCCTTT--GTCCAGCTGTCCCATTCAGTTCAGACTGCTCCC 4477
 Qy 2461 GGGGAGAGCCCGCCCGCCCGCAGACATAAGAACTGACGCTTGGTACTGACAGAGTCTG 2520
 Db 4478 GGGGAGAGCCCGCCCGCCCGCAGACATAAGAACTGACGCTTGGTACTGACAGTCTG 4537
 Qy 2521 GGTGTGTAGAGAACTCTTTGTAAAGCAATAAGTTTGGGTGTATGA CAATGTTAAAAA 2579
 Db 4538 GGTGTGTAGAGAACTCTTTGTAAAGCAATAAGTTTGGGTGTATGA CAATGTTAAAAA 4596

QY 1261 ACAAAACAGGGATCTTTTTCGCCCCCTGCTCCAGTNCAGAGTTGGCTGNACCCGCTTGG 1320
Db |||||
QY 3341 ACAAAACAGGGATCTTTTTCGCCCCCTGCTCCAGT- CGAGTTGGCTGA--CCGCTTG 3396
Db |||||
QY 1321 ANTCACTGACCAATTTGTTGGAGANACAGGGAGAGAGCTTCCAGCCTGGGTGAGAAAGG 1380
Db |||||
QY 3397 GATCAGTGACCAATTTGTTGGCAGA-CAGGGAGAGAGAGCTTCCAGCCTGGGTGAGAAAGG 3455
Db |||||
QY 1381 GTGGCGAGCCCTTCGGGCCCTCACCCCTNCCAGGCTGCTGTGNAGAGTCAAGTGTGA 1440
Db |||||
QY 3456 GTGGCGAGCCCTTCGGGCCCTTCACCCCT-CGAGGCTGCTGTG-AGAGTGTCAAGTGTGA 3513
Db |||||
QY 1441 AGGNGCCCAAAANCTCAGGNTTCAGTGCAGAACACAGGTNCCAGAGTATGCCCGCCGNTA 1500
Db |||||
QY 3514 AGGGCCCAAACTC--AGGTTTCAGTGCAGAACAGGT-CAGCAGGTATGCCCGCCGCTAG 3569
Db |||||
QY 1501 GGTTAANNNGGGGCGCTCTNAAACCCCTTGGCCCTNCCAGTTCACCTNCCAGCTCAGCTC 1560
Db |||||
QY 3570 GTTAA---GGGGGCCCTCTAAACCCCTTGGCTGGCCCTCA---CCTGGCCAGCTCA-CCC 3621
Db |||||
QY 1561 CTTTGTGGGTAGGGGAAAGAAATGCTGACCCCTGGGAAGGCTWCCCTGGTAGATACAC 1620
Db |||||
QY 3622 CTTTGTGGGTAGGGGAAAGAAATGCTGACCCCTGGGAAGGCT-CCCTGGTAGATACAC 3680
Db |||||
QY 1621 CACACTTTTCAGGTTGTGCAACACAGGTCTCTGAGTTGACCTCTGTTTCAGCCAAAGGACC 1680
Db |||||
QY 3681 CACACTTTTCAGGTTGTGCAACACAGGTCTCTGAGTTGACCTCTGTTTCAGCCAAAGGACC 3740
Db |||||
QY 1681 AAAGAAGGTGTAAAGTGAAGTGTTCAGTNCNCCAGACATGTGCCCTTTGCTGTG 1740
Db |||||
QY 3741 AAAGAAGGTGTAAAGTGAAGTGTTCAGT-CCCAAGACATGTGCCCTTTGCTGTG 3799
Db |||||
QY 1741 GCTACCACTCTTCCAGAGCAGCAGGCCCCGAGCCCTTCAGGCCAGCACTGCCCCAG 1800
Db |||||
QY 3800 GCTACCACTCTTCCAGAGCAGCAGGCCCCGAGCCCTTCAGGCCAGCACTGCCCCAG 3859
Db |||||
QY 1801 ACTCGCTGGCACTCAGTTTCCCTCATCTGTAAAGGTGAAGGGGTGATGCAGGATATCCCTGA 1860
Db |||||
QY 3860 ACTCGCTGGCACTCAGTTTCCCTCATCTGTAAAGGTGAAGGGGTGATGCAGGATATCCCTGA 3919
Db |||||
QY 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAGNAAAGCAGCAGAGAGAGAGY 1920
Db |||||
QY 3920 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAA--GGAAAGCAGCAGAGAGAGAGC- 3976
Db |||||
QY 1921 TCCGGGCCCCCAGNCCCCACTNATCAGTGTNCCAGCGTCTNGTTCNCCAGNAGCACA 1980
Db |||||
QY 3977 TCCGGGCCCCCAGNCCCCACTATCAGTGT---CCAGCGTGTGGTTCCCCAGAGCAGC- 4032
Db |||||
QY 1981 GCTNCAGNCACTCAGCACTNCACTNCCCTGCTGCCCCCTGCCCCGANGAGGGTACTG 2040
Db |||||
QY 4033 -----TCAGATCACACTGACATCACTACCTGCCCCCTGCCCCCTGG--CCAGGGGTACTG 4083
Db |||||
QY 2041 CCGNACGGCACTTCACNTCTGATGNAACCTCAAGACATTTTCATGCTGCTGCTTNG 2100
Db |||||
QY 4084 CCG-ACGGCACTTTGCACT-CTGATGACCTCAAAAGCACTTTTCATGCTGCTGCTTGGCA 4140
Db |||||
QY 2101 GCAGGNCAGGNCAGGNCAGTGCACNCTGTAGNAGCATANGCAANGCCAGGATGG 2160
Db |||||
QY 4141 GGGCAGGGCAGGGCAGTGACACT-----GTAGGAGCATATGCAAGCCAGGAGATGG 4190
Db |||||
QY 2161 GGTGNAAGGAGNACAGTCTTTCAGCTGTCCANCACTGTCGACTNCTCAAAACCTCTTN 2220
Db |||||
QY 4191 GGTGAA--GGGACACAGTCTTTCAGCTGTCCA-CATGCACTGACT-CTCAAACTCTT- 4245
Db |||||
QY 2221 NCCAGNATTTCTTAAGATAGCAGNCCCTTNNCCCAATGCCCCAGCTTAGCTCTTCT 2280
Db |||||
QY 4246 --CCAGATTTCTTAAGATAGCAGNCCCTTNNCCCAATGCCCCAGCTTAGCTCTTCT 4301
Db |||||
QY 2281 CCCAGGGAGCTTANTCAGGACTCAGTAGCAATTAATCAGCTGTGNAATCGTCAGGGG 2340
Db |||||
QY 4302 CCCAGGGAGCTA-CTCAGGACTCAGTAGCAATTAATCAGCTGTG-AATCGTCAGGGG 4359

QY 2341 TGTCTGTAGCTCAACCTCTCTGGGGCAGGGGACGCCGAGACTCCTGTTGGGAGAGCTCAT 2400
Db |||||
QY 4360 TGTCTGTAGCTCAACCTCTCTGGGGCAGGGGACGCCGAGACTCCTGTTGGGAGAGCTCAT 4419
Db |||||
QY 2401 TCCACATCTTGGCCAGACAGCCTTTNGTCCAGCTGTCCACATTCAGTCAGACTGCTCCC 2460
Db |||||
QY 4420 TCCACATCTTGGCCAGACAGCCTTT-GTCCAGCTGTCCACATTCAGTCAGACTGCTCCC 4478
Db |||||
QY 2461 GGGGAGAGAGCCCGGGCCCCCAGCACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTG 2520
Db |||||
QY 4479 GGGGAGAGAGCCCGGGCCCCCAGCACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTG 4538
Db |||||
QY 2521 GGTGTAGAGAACTCTTTTGTAAAGCAATAAAGTTGGGGTGTATGACAAATGTTAAAAA 2578
Db |||||
QY 4539 GGTGTAGAGAACTCTTTTGTAAAGCAATAAAGTTGGGGTGTATGACAAATGTTAAAAA 4596
Db |||||
RESULT 4
ABK51171
ID ABK51171 standard; cDNA; 4596 BP.
XX
AC ABK51171;
XX
DT 30-JUL-2002 (first entry)
XX
DE cDNA encoding human cellular kinase NIK protein.
XX
KW Human; virucide; cytomegalovirus infection; CMV; cellular kinase; RICK;
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 233..3076
FT /*tag= a
FT /product= "Human cellular kinase NIK (Nck-Interacting
FT kinase)"
XX
EP1201765-A2.
XX
PD 02-MAY-2002.
XX
PF 15-OCT-2001; 2001EP-00124604.
XX
PR 16-OCT-2000; 2000US-0240750P.
XX
PA (AXXI-) AXXIMA PHARM AG.
XX
PI Schubart D, Habenberger P, Stein-Gerlach M, Bevec D;
XX
DR WPI; 2002-373930/41.
XX
DR P-PSDB; AAU80371.
XX
PT Identifying agents for treatment or prevention of cytomegalovirus
PT infection, comprises contacting test compound with cellular kinase and
PT detecting change in cellular kinase activity.
XX
PS Disclosure; Page 29-34; 49pp; English.
XX
CC The present invention relates to a new method for identifying compounds
CC for treating and/or preventing cytomegalovirus (CMV) infection and/or
CC related diseases. The method of the invention comprises contacting a test
CC compound with at least one of the cellular kinases RICK, RIP, Nck-
CC Interacting kinase, MKK3 and SRPK-2 and detecting any change in kinase
CC activity. The method of the invention can be used to treat and/or prevent
CC CMV infections and related diseases. Oligonucleotides that can detect the
CC specified kinases can also be used for diagnosis of infection. The
CC present nucleic acid sequence encodes the human cellular kinase NIK (Nck-
CC Interacting kinase) protein of the invention, as described above
XX
SQ Sequence 4596 BP; 1055 A; 1401 C; 1299 G; 841 T; 0 U; 0 Other;
Query Match 70.3%; Score 1850.2; DB 6; Length 4596;

Best Local Similarity 90.5%; Pred. No. 0;
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

Qy 1 CCCCTCTCACAGCCAGCCCATCCAAAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT 60
Db CCCCTCTCACAGCCAGCCCATCCAAAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT 2159
Qy 61 CTGACAGCGAGCTGGAGGAGAGGTGAACCGGCACTACAGCAAGTGGAGGTCTGAAGA 120
Db CTGACAGCGAGCTGGAGGAGAGGTGAACCGGCACTACAGCAAGTGGAGGTCTGAAGA 2219
Qy 121 GCCCTTGGAGGGAGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 180
Db GCCCTTGGAGGGAGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 2279
Qy 181 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCGAGGGCCCGGCCAG 240
Db ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCGAGGGCCCGGCCAG 2339
Qy 241 CTGAGGAGACAACAGGCGAGAGCCCTTAAGTCTCAGCCTCTCTCCACACAGAGCCCGCAG 300
Db CTGAGGAGACAACAGGCGAGAGCCCTTAAGTCTCAGCCTCTCTCCACACAGAGCCCGCAG 2399
Qy 301 AGCCAAACAAGTCTCCCTTGAAGTCTTGAAGAGGAGGAGTCTGGAGTGGGAAACCT 360
Db AGCCAAACAAGTCTCCCTTGAAGTCTTGAAGAGGAGGAGTCTGGAGTGGGAAACCT 2459
Qy 361 TACCTCTGTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACCAGAGCGGAAAG 420
Db TACCTCTGTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACCAGAGCGGAAAG 2519
Qy 421 CAACCGTCCCGAGCAGGAACTGACAGAGCTGGAATAGAAATTAATCTCAACAGCCTGT 480
Db CAACCGTCCCGAGCAGGAACTGACAGAGCTGGAATAGAAATTAATCTCAACAGCCTGT 2579
Qy 481 CCGAGCCATTTCTCTGGAGAGCAGAGGCAAAATCTCTGTGCTCAGCATCAGACGCC 540
Db CCGAGCCATTTCTCTGGAGAGCAGAGGCAAAATCTCTGTGCTCAGCATCAGACGCC 2639
Qy 541 TCTCCCTGTCCGATGACAGTGAAGAAACCCATCAAAAGGCTCTCAAAAGCTCGCGGACA 600
Db TCTCCCTGTCCGATGACAGTGAAGAAACCCATCAAAAGGCTCTCAAAAGCTCGCGGACA 2699
Qy 601 CCCTGAGCTCAGGCGTACACTCTGGAGCAGCCAGGCGGAGCTCGAAGCTTCCAGCTGGA 660
Db CCCTGAGCTCAGGCGTACACTCTGGAGCAGCCAGGCGGAGCTCGAAGCTTCCAGCTGGA 2759
Qy 661 ACATGGTGTGGCCGGGGGGGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 720
Db ACATGGTGTGGCCGGGGGGGGGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 2819
Qy 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCGGGAGTTCACCGGTCAAG 780
Db TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCGGGAGTTCACCGGTCAAG 2879
Qy 781 TGGAGACATCGCCACCTGGCATCAGCAGCCAGATCCAGCTTCAAGCTTGGTCA 840
Db TGGAGACATCGCCACCTGGCATCAGCAGCCAGATCCAGCTTCAAGCTTGGTCA 2939
Qy 841 CCAAAGACGGGACCTGTTCCGTACGACATGAGAGTGCAGATCCGGGATCGACCTGC 900
Db CCAAAGACGGGACCTGTTCCGTACGACATGAGAGTGCAGATCCGGGATCGACCTGC 2999
Qy 901 AGTGCACACTGGCCCTGTAGGAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 960
Db AGTGCACACTGGCCCTGTAGGAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 3059
Qy 961 TGGAGAACAGGCCCTTAACCTTGCCTCCACCGCGGCTCCACACTGCGGAAAGAGCCT 1020
Db TGGAGAACAGGCCCTTAACCTTGCCTCCACCGCGGCTCCACACTGCGGAAAGAGCCT 3119
Qy 1021 TCCTGCTCGGTGACAGTGTGCCCTGAAACACAGAGCTCAGCCGTTCCAGGGGATYTG 1080
Db TCCTGCTCGGTGACAGTGTGCCCTGAAACACAGAGCTCAGCCGTTCCAGGGGATYTG

Db 3120 TCCTGCTCGGTGACAGTGTGCCCTGAAACACAGAGCTCAGCCGTTCCACAGGGGATTT-- 3177
Qy 1081 NCCAGCCCCCGGCTCARCAGNTGGGAACACAGGCGCTCGNCAGCAGVAGNAAGTNGGGGG 1140
Db GCGAGCCCCCGGCTCA--CAGTGGAAACAGGSCCTCG-----CAGCAGCAAGTGGGG 3230
Qy 1141 CAGCNAGATGCTCTCCAGGATTTCAACNCTGAGCCNTGCCANCCCTGCTGAADA 1200
Db GCAAGCAGAAATGCTCTCCAGGATTTCACTGAGCCCTGCCAC-----CCTGCTGA 3284
Qy 1201 AAACAYTNCCGCCACGTGAAGAGACAGAGGATGNCAGGAGTTNNACCTTGGGGAA 1260
Db AAACATCCGCCACGTGAAGAGACAGAGGATGNC-----AGGATTTACCTTGGGAA 3340
Qy 1261 AAAAAACAGGGATCTTTTCTGCCCTGTCTCAGTNCAGATTTGGCCTGNAACCGCTTGG 1320
Db AAAAAACAGGGATCTTTTCTGCCCTGTCTCAGTNCAGATTTGGCCTGNAACCGCTTGG 3396
Qy 1321 ANTCACTGACATTTGTTGGCAGANAGGGGAGAGAGCTTTCAGCCTGGGTCAAGAGGG 1380
Db GATCAGTGACATTTGTTGGCAGACAGGGGAGAGAGCTTTCAGCCTGGGTCAAGAGGG 3455
Qy 1381 GTGGCGAGCCCTTGGSCCTCACCCTNCCAGGCTCTGTGNAGAGTGTCAAGTGTGTA 1440
Db GTGGCGAGCCCTTGGSCCTTACCCT--CAGGCTGCTGTG--AGAGTGTCAAGTGTGTA 3513
Qy 1441 AGGNCNCCAAANCTCAGGNTTCAGTGCAGAACACAGGTNCCAGCAGTATGCCCCCGNNTA 1500
Db AGGNCNCCAAATC--AGGTTTCAGTGCAGAACACAGGT--CAGCAGTATGCCCCCGTAG 3569
Qy 1501 GGTAAANNNGGGGCGCTCTNAAACCCCTTGGCTNCCAGCTTCCAGCTCAGCTCAGCCT 1560
Db GTTAA-----GGGGCGCTCTTAAACCCCTTGGCTGCGCTCA---CCTGGCCAGCTCA--CCC 3621
Qy 1561 CTTTGGGTGAGGGGAAAGATGCTGACCTTGGGAAGCTTCCCTGCTGAGTATACAC 1620
Db CTTTGGGTGAGGGGAAAGATGCTGACCTTGGGAAGCTTCCCTGCTGAGTATACAC 3680
Qy 1621 CACACTTTTCAGGTTGTGCAACACAGGCTCTGAGTTGACCTCTGGTTACGCCAAGACC 1680
Db CACACTTTTCAGGTTGTGCAACACAGGCTCTGAGTTGACCTCTGGTTACGCCAAGACC 3740
Qy 1681 AAAGAGGTGTAAAGTGAAGTGTCTCAGTNCNCCAGACATGTGCCCCCTTTCTGCTG 1740
Db AAAGAGGTGTAAAGTGAAGTGTCTCAGT--CCCCAGACATGTGCCCCCTTTCTGCTG 3799
Qy 1741 GCTACCACTCTTCCAGAGCAGCGCCCGAGCCCTTTCAGGCCAGCACTGCCCCAG 1800
Db GCTACCACTCTTCCAGAGCAGCGCCCGAGCCCTTTCAGGCCAGCACTGCCCCAG 3859
Qy 1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAGTGAAGGTGATGAGGATATGCTGTA 1860
Db ACTCGCTGGCACTCAGTTCCCTCATCTGTAAGTGAAGGTGATGAGGATATGCTGTA 3919
Qy 1861 CAGGAAACAGTCTGTGGATGACATGATCAGTGTCTTAAAGNAAAGCAGCAGAGAGACGY 1920
Db CAGGAAACAGTCTGTGGATGACATGATCAGTGTCTTAAAGNAAAGCAGCAGAGAGACG 3976
Qy 1921 TCCGGCCCCCAGNCCCACTNATCAGTGTNCCAGCGTGTCTTCCAGGAGCAGCA 1980
Db TCCGGCCCCCAGNCCCACTNATCAGTGTNCCAGCGTGTCTTCCAGGAGCAGCA-- 4032
Qy 1981 GCTNAGNCACTCANCACTNACCTNCCCTGCGCTTCCCTTNGGCCANAGGGTACTG 2040
Db -----TCAGCATCACTGACACTCACCCTGCCCTGCCCTGG--CCAGAGGGTACTG 4083
Qy 2041 CCGNACGGCACTTTGCACTNCTGATGNACCTCAAGCACTTTTCATGCTNGCCCTCTNNG 2100
Db CCG-ACGGCACTTTGCACT--CTGATGACCTCAAGACACTTTTCATGCTGCCCTCTGGCA 4140
Qy 2101 GCAGGNCAGGNCAGGNCAGTGCACNCTGTAGGNCATANGCAAGCCAGGAGATGG 2160
Db GGCAGGGCAGGCGAGTGACACT-----GTAGGAGCATAGCAAGCCAGGAGATGG 4190

QY 2161 GGTGNAAGGACACAGTCTTGAGCTGTCCANATGATGCTGACCTNCTCAAACTCTTN 2220
 |||||
 Db 4191 GGTGAA--GGGACACAGTCTTGAGCTGTCCA-CATGCACTGACT-CTCAAACTCTTT- 4245
 |||||
 QY 2221 NCCAGNATTTCTTAAGATAGCANCCCTTTNCCCATTTGCCAGCTTAGCTCTTCT 2280
 |||||
 Db 4246 --CCGATTTCTTAAGATAGCACCCCC--TTCCCATTTGCCAGCTTAGCTCTTCT 4301
 |||||
 QY 2281 CCCAGGGAGCTTANCTCAGGACTCAGTAGCATTAATCAGCTGTGNAATCTGAGGGG 2340
 |||||
 Db 4302 CCCAGGGAGCTA-CTCAGGACTCAGTAGCATTAATCAGCTGTG-AATCGTCAAGGGG 4359
 |||||
 QY 2341 TGTCTGCTAGCTCAACCTTCCTGGGCGAGGGACGCCGAGACTCCGTTGGGAGAACTCAT 2400
 |||||
 Db 4360 TGTCTGCTAGCTCAACCTTCCTGGGCGAGGGAGCGCCGAGACTCCGTTGGGAGAACTCAT 4419
 |||||
 QY 2401 TCCACATCTTCCCAAGACAGCCTTTNFTCCAGCTGTCCACATTCAGTTCAGCTCTCC 2460
 |||||
 Db 4420 TCCACATCTTCCCAAGACAGCCTTTT-GTCCAGCTGTCCACATTCAGTTCAGCTCTCC 4478
 |||||
 QY 2461 GGGGAGAGACCCCGGCCCGCCAGCATATAAGAACTGCGACTTGGTACTGCGAGTCTG 2520
 |||||
 Db 4479 GGGGAGAGACCCCGGCCCGCCAGCATATAAGAACTGCGACTTGGTACTGCGAGTCTG 4538
 |||||
 QY 2521 GGTCTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGTATGACAAATGTTAAAAA 2578
 |||||
 Db 4539 GGTCTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGTATGACAAATGTTAAAAA 4596
 |||||

RESULT 5
 ADE85075
 ID ADE85075 standard; DNA; 4596 BP.
 XX AC ADE85075;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Farnesyl transferase inhibitor modulated leukemia associated gene #294.
 XX
 KW ss; cytostatic; farnesyl transferase inhibitor; gene expression;
 KW quinolone; leukemia; cancer.
 OS Homo sapiens.
 XX
 PN WO2003038129-A2.
 XX
 PD 08-MAY-2003.

30-OCT-2002; 2002WO-US034784.
 PF
 XX
 PR 30-OCT-2001; 2001US-0338997P.
 PR 30-OCT-2001; 2001US-0340081P.
 PR 30-OCT-2001; 2001US-0340938P.
 PR 30-OCT-2001; 2001US-0341012P.
 XX
 (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
 PA
 XX
 PI Raponi M;
 XX
 DR WPI; 2003-513497/48.
 XX
 PT Determining whether a patient will respond to treatment with a farnesyl
 PT transferase inhibitor, by analyzing the expression of gene that is
 PT differentially modulated in the presence of the inhibitor.
 XX
 PS Disclosure; SEQ ID NO 294; 346pp; English.
 XX

The invention relates to a method of determining whether a patient will
 CC respond to treatment with a farnesyl transferase inhibitor (FTI), by
 CC analyzing the expression of gene that is differentially modulated in the
 CC presence of an FTI. The method is useful for determining whether a
 CC patient will respond to treatment with a FTI such as (B)-6-(amino(4-

CC chlorophenyl) (1-methyl-1H-imidazol-5-yl)methyl-4-(3-chlorophenyl)-1-
 CC methyl-2-(1H)quinoline, monitoring the therapy of a patient, treating a
 CC patient with leukemia with FTI if the analysis indicates that the patient
 CC will respond. This sequence corresponds to a gene whose expression may be
 CC modulated in the presence of FTI.

SQ Sequence 4596 BP; 1055 A; 1401 C; 1299 G; 841 T; 0 U; 0 Other;

Query Match 70.3%; Score 1850.2; DB 10; Length 4596;
 Best Local Similarity 90.5%; Pred. No. 0;
 Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

QY 1 CCCCTCTCACAGCCAGGCCCATCAAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT 60
 Db 2100 CCCCTCTCACAGCCAGGCCCATCAAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT 2159
 |||||
 QY 61 CTGACGGAGCTGGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA 120
 Db 2160 CTGACGGAGCTGGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA 2219
 |||||
 QY 121 GCCCTTGGAGGGGAGATATATAAGAACCAAGACATCAACCGCCAAATCAAGCCAAATACC 180
 Db 2220 GCCCTTGGAGGGGAGATATATAAGAACCAAGACATCAACCGCCAAATCAAGCCAAATACC 2279
 |||||
 QY 181 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCCGCAAGGGGCCCGAGGGCCCGGCCAG 240
 Db 2280 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCCGCAAGGGGCCCGAGGGCCCGGCCAG 2339
 |||||
 QY 241 CTGAGGAGACAAACAGGAGAGCCCTTAAGTCTCAGCTCTCTCTCCACAGAGCCCGCCAG 300
 Db 2340 CTGAGGAGACAAACAGGAGAGCCCTTAAGTCTCAGCTCTCTCTCCACAGAGCCCGCCAG 2399
 |||||
 QY 301 AGCCAAACAGTCT 360
 Db 2400 AGCCAAACAGTCT 2459
 |||||
 QY 361 TACCTCTGTCT 420
 Db 2460 TACCTCTGTCT 2519
 |||||
 QY 421 CAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAAATAGAAATATTCTCTCAACAGCTGT 480
 Db 2520 CAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAAATAGAAATATTCTCTCAACAGCTGT 2579
 |||||
 QY 481 CCCAGCATTTCT 540
 Db 2580 CCCAGCATTTCT 2639
 |||||
 QY 541 TCTCCCTGTGGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGGCTCGCGGGACA 600
 Db 2640 TCTCCCTGTGGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGGCTCGCGGGACA 2699
 |||||
 QY 601 CCCTGAGCTCAGCGGTACATCTCTGGAGCAGCCAGCCGAGGCTCGAAGCTCCAGCTGGA 660
 Db 2700 CCCTGAGCTCAGCGGTACATCTCTGGAGCAGCCAGCCGAGGCTCGAAGCTCCAGCTGGA 2759
 |||||
 QY 661 ACATGGTCTGGCCCGGGGGGGCCACCCAGCACCCCAAGCTATTTCATGGTGTGAAG 720
 Db 2760 ACATGGTCTGGCCCGGGGGGGCCACCCAGCACCCCAAGCTATTTCATGGTGTGAAG 2819
 |||||
 QY 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCCACCGGGTCAAG 780
 Db 2820 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCCACCGGGTCAAG 2879
 |||||
 QY 781 TGGGAGACATCGGCATCTGGCATCAGCAGCAGATCCAGCTGAGCTTTCAGCTTGTGCA 840
 Db 2880 TGGGAGACATCGGCATCTGGCATCAGCAGCAGATCCAGCTGAGCTTTCAGCTTGTGCA 2939
 |||||
 QY 841 CCAAGAGCGGCGCTCTTCT 900
 Db 2940 CCAAGAGCGGCGCTCTTCT 2999
 |||||
 QY 901 AGTGACACTGGGCCCTCTGTATGGGAGCTTTCGCTGGAGGTCAAGCATGGCCAGC 960

XX The invention relates to a method of classifying a breast cancer patient
CC according to prognosis by determining the similarity between the level of
CC expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
CC control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
CC marker used in the method of the invention.

XX SQ Sequence 4596 BP; 1055 A; 1401 C; 1299 G; 841 T; 0 U; 0 Other;

Query Match 70.3%; Score 1850.2; DB 13; Length 4596;

Best Local Similarity 90.5%; Pred. No. 0; Mismatches 159; Indels 81; Gaps 34;

Matches 2334; Conservative 4;

QY 1 CCCCTCTCACAGCCAGGCCATCCAAAGAGGGCTGAGGAAGAGCCCATCCACCGCGTGT 60
DB 2100 CCCCTCTCACAGCCAGGCCATCCAAAGAGGGCTGAGGAAGAGCCCATCCACCGCGTGT 2159

QY 61 CTGACGCGGAGCTGGGAGGAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120
DB 2160 CTGACGCGGAGCTGGGAGGAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 2219

QY 121 GCCCTTGGAGGGGAGATATAAGAACCAAGACATCCACGCCAAATCAAGCCAAATTACC 180
DB 2220 GCCCTTGGAGGGGAGATATAAGAACCAAGACATCCACGCCAAATCAAGCCAAATTACC 2279

QY 181 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGAGGCCCGGCCAG 240
DB 2280 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGAGGCCCGGCCAG 2339

QY 241 CTGAGAGACAACAGGAGAGCCCTAAGCTCAGAGCTCTCTCCACAGAGCCCGCCAG 300
DB 2340 CTGAGAGACAACAGGAGAGCCCTAAGCTCAGAGCTCTCTCCACAGAGCCCGCCAG 2399

QY 301 AGCCAAACAAGTCTCCTCCCTTGTACTTTTGAGCAAGGAGAGTCTGGGATGTGGGAACCT 360
DB 2400 AGCCAAACAAGTCTCCTCCCTTGTACTTTTGAGCAAGGAGAGTCTGGGATGTGGGAACCT 2459

QY 361 TACCTCTGTCTCCTTGGAGCCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAG 420
DB 2460 TACCTCTGTCTCCTTGGAGCCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAG 2519

QY 421 CAACCGTCCGAGCAGGAACTGCAGCAGCTGGAATAGAAATATTCCTCAACAGCCTGT 480
DB 2520 CAACCGTCCGAGCAGGAACTGCAGCAGCTGGAATAGAAATATTCCTCAACAGCCTGT 2579

QY 481 CCCAGCATTTTCTCTGGAGGAGCAGGAGCAAAATCTCTCGTGCCTCAGCATCGACAGCC 540
DB 2580 CCCAGCATTTTCTCTGGAGGAGCAGGAGCAAAATCTCTCGTGCCTCAGCATCGACAGCC 2639

QY 541 TCTCCTCTGTCCGATGACAGTGAAGAACCCATCAAAAGCCCTCTCAAAGCTCCGCGGACA 600
DB 2640 TCTCCTCTGTCCGATGACAGTGAAGAACCCATCAAAAGCCCTCTCAAAGCTCCGCGGACA 2699

QY 601 CCTTGAGCTCAGGCTGACACTCTCTGAGCAGCAGCCGAGGCTCGAAGCTCAGCTGGA 660
DB 2700 CCTTGAGCTCAGGCTGACACTCTCTGAGCAGCAGCCGAGGCTCGAAGCTCAGCTGGA 2759

QY 661 ACATGTGTCTGGCCGGGGCGGCCCAACGACACCCCAAGCTATTTCAATGGTGTGAAG 720
DB 2760 ACATGTGTCTGGCCGGGGCGGCCCAACGACACCCCAAGCTATTTCAATGGTGTGAAG 2819

QY 721 TCCAAATACAGTCTCTTAATGTGAACA CTTGCACATCCGGGAGTTTCCA CCGGGTCAAG 780
DB 2820 TCCAAATACAGTCTCTTAATGTGAACA CTTGCACATCCGGGAGTTTCCA CCGGGTCAAG 2879

QY 781 TGGGAGACATCGCACTGGCATCAGCAGCAGATCCAGCTGCAGCCTTCACTTGGTCA 840
DB 2880 TGGGAGACATCGCACTGGCATCAGCAGCAGATCCAGCTGCAGCCTTCACTTGGTCA 2939

QY 841 CCAAAGACGGGACAGCCTGTTCGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC 900
DB 2940 CCAAAGACGGGACAGCCTGTTCGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC 2999

QY 901 AGTGCACACTGGCCCTCGATGAGGAGCTTCGCTGGAGCTGGAGGTTCAAGCATGGCCAGC 960
DB 3000 AGTGCACACTGGCCCTCGATGAGGAGCTTCGCTGGAGCTGGAGGTTCAAGCATGGCCAGC 3059

QY 961 TGGAGAACAGCCCTAAACCTCCACCGCGGCTCCACACTGCGGAAAGCAGGCT 1020
DB 3060 TGGAGAACAGCCCTAAACCTCCACCGCGGCTCCACACTGCGGAAAGCAGGCT 3119

QY 1021 TCCTGCTCGGTGCACGATGCTGCCCTGAAAAACACAGGCTCAGCGCTTCCCAGGGGATYTG 1080
DB 3120 TCCTGCTCGGTGCACGATGCTGCCCTGAAAAACACAGGCTCAGCGCTTCCCAGGGGATYTG 3177

QY 1081 NCGAGCCCCCGGCTCARGAGTGGGAACAGGCGCTCGNACAGNAGNAGGTTNGGGG 1140
DB 3178 GCCAGCCCCCGGCTCA--CAGTGGGAACACAGGCGCTCG-----CAGCAGCAAGGTGGGG 3230

QY 1141 CAAGCNAGAAATGCTCCAGGATTTTCAACNCTGAGCCCTGCCCCANCCCTGCTGAADA 1200
DB 3231 GCAAGCAGAAATGCTCCAGGATTTTCAACNCTGAGCCCTGCCCCAC-----CCTGCTGA 3284

QY 1201 AAACAYTNCCGCCACGTGAAGAGAGAGAGGATGNCAGAGTNNACTTGGGGAA 1260
DB 3285 AAAAACATCCGCCACGTGAAGAGAGAGAGGATGNC-----AGGAGTTTACTTGGGAA 3340

QY 1261 ACAAAAAGGAGATCTTTTCTGCGCCCTGCTCAGTNCAGTGGCTGNCACCGCTTGG 1320
DB 3341 ACAAAAAGGAGATCTTT--TTCTGCGCCCTGCTTCCAGT--CGAGTTGGCCCTGA--CCCGCTTG 3396

QY 1321 ANTCACTGACCATTTGTTGGCAGANACAGGGGAGAGAGCTTCCAGCCTGGGTGAGAAAGG 1380
DB 3397 GATCAGTACCATTTGTTGGCAGA--CAGGGAGAGAGCTTCCAGCCTGGGTGAGAAAGG 3455

QY 1381 GTGGGAGAGCCTTTCGGCCCTCAACCTNCCAGGCTGCTGTGNAAGAGTCAAGTGTGTA 1440
DB 3456 GTGGGAGAGCCTTTCGGCCCTCAACCT--CCAGGCTGCTGTG--AGAGTGTCAAGTGTGTA 3513

QY 1441 AGGNCACAAANTCAGGTTTCAAGTGCAGAACAGGTTNACAGAGTATGCCCGCCGNTA 1500
DB 3514 AGGNCACAAATC---AGGTTTCAAGTGCAGAACAGGTT--CAGAGGATATGCCCGCCGNTAG 3569

QY 1501 GCTTAANNNGGGGGCCCTCTNAAAACCCCTTGCTNGGCCCTNCACTNGGCCAGCTCANCC 1560
DB 3570 GTTAA----GGGGGCCCTCTAAAACCCCTTGCTGCGCTCA--CCTGGCCAGCTCA--CCC 3621

QY 1561 CTTTTGGGTGTAGGGGAAAAAGATGCTTGA CTTGGGAAGGCTWCCCTGGTAGAATACAC 1620
DB 3622 CTTTTGGGTGTAGGGGAAAAAGATGCTTGA CTTGGGAAGGCT--CCCTGGTAGAATACAC 3680

QY 1621 CACACTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCTGCTTTCAGCAGGAGCC 1680
DB 3681 CACACTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCTGCTTTCAGCAGGAGCC 3740

QY 1681 AAAGAAAGTGTGAAGTGAAGTGTCTCAGTNCCTCCAGACATGTCGCCCTTTTGTGCTG 1740
DB 3741 AAAGAAAGTGTGAAGTGAAGTGTCTCAGT--CCCCAGACATGTCGCCCTTTTGTGCTG 3799

QY 1741 GCTACCACTTTTCCCAGAGCAGGCCCCCGAGCCCTTACAGGCCACGACTGCCCCAG 1800
DB 3800 GCTACCACTTTTCCCAGAGCAGGCCCCCGAGCCCTTACAGGCCACGACTGCCCCAG 3859

QY 1801 ACTCGCTGGCACTCAGTTTCCCTCATCTGTAAGGTGAAGGGTGTATGCAAGGATATGCTGA 1860
DB 3860 ACTCGCTGGCACTCAGTTTCCCTCATCTGTAAGGTGAAGGGTGTATGCAAGGATATGCTGA 3919

QY 1861 CAGGAACAGTCTGTGGATGGAATGATCAGTGTCTNAGGNAAGCAGCAGAGAGAGCY 1920
DB 3920 CAGGAACAGTCTGTGGATGGAATGATCAGTGTCTNAGGNAAGCAGCAGAGAGAGAGCY--GGAAAGCAGCAGAGAGAGAGCG-- 3976

1921	TCGGCGCCCGCAGNCCCACTNATCAGTGTGNCAGCGTGTCTGCTGTTNCCCGCAGNAGCAC	1980
Qy		
3977	TCGGCGCCCGCAGNCCCACTATCAGTGT-----CCAGCGTGTGTGTTCCCGCAGAGCACAGC-	4032
Db		
1981	GCTNCAGNCATCANCACTGCACATNCACCTGTCGCTGCGCTTGGCCGTCGAGAGGGTACTG	2040
Qy		
4033	-----TCAGCATCACACTGCACACTCACCTGCGCTGCGCCCTGG--CGAGAGGGTACTG	4083
Db		
2041	CCGNAGGCGACTTTGCACNCTCTGATGNACCTCAAAGCATTTTCATGGCTGNGCCCTCTTNG	2100
Qy		
4084	CCG-ACGGCACTTTGCACT--CTGATGACCTCAAAGCACATTTTCATGGCTGCCCTCGGCA	4140
Db		
2101	GCAGGNCAGGNCAGGNCAGGNCAGTGCACANCTGTAGGNAGCATANGCANGCCAGGAGATGG	2160
Qy		
4141	GGGCGAGGCGAGGGCAGTGCACACT-----GTAGGAGCATAGCAAGCCAGGAGATGG	4190
Db		
2161	GGTGNAGGGANCAACAGTCTTTGAGCTGTCCANCATGATGTGACTNCTCAAACCTCTTN	2220
Qy		
4191	GGTGAA--GGGACACAGTCTTTGAGCTGTCCA-CATGCAATGTGACT--CTTCAAACCTCTT-	4245
Db		
2221	NCAGNATTTCTCTAAGAAATAGCANCCCCCTTNCGCCATTTGCCCGCAGCTTAGCCTCTTCT	2280
Qy		
4246	--CCAGATTTCTCTAAGAAATAGCACCCCC--TTCCCCATTTGCCCGCAGCTTAGCCTCTTCT	4301
Db		
2281	CCAGGGGAGCTANCTCAGGACTCACGTAGCATTTAAATCAGCTGTGNAATCGTCAGGGGG	2340
Qy		
4302	CCAGGGGAGCTA-CTCAGGACTCACGTAGCATTTAAATCAGCTGTG--ATTCGTCAAGGGG	4359
Db		
2341	TGTCGTGAGCTCTCAACTCTCGGGCAGGGGAGCGCGAGACTCCGTGGGAGAGCTCAT	2400
Qy		
4360	TGTCGTAGCTCTCAACTCTCGGGCAGGGGAGCGCGAGACTCCGTGGGAGAGCTCAT	4419
Db		
2401	TCCCATCATTTGCCAAGACAGCCTTTNGTCCAGCTGTCCACATTTAGTCAGACTGTCTCCC	2460
Qy		
4420	TCCCATCATCTTGGCAAGACAGCCTTTT--GTCCAGCTGTCCACATTTGAGTCAGACTGTCTCCC	4478
Db		
2461	GGGAGAGAGCCCGGGCCCCCAGCACATAAAGAACTGCAGCCTTTGGTACTGCAGAGTCTG	2520
Qy		
4479	GGGAGAGAGCCCGGGCCCCCAGCACATAAAGAACTGCAGCCTTTGGTACTGCAGAGTCTG	4538
Db		
2521	GGTTGTAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGTATGACAAATGTTAAAAA	2578
Qy		
4539	GGTTGTAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGTATGACAAATGTTAAAAA	4596
Db		

RESULT 7

ACN44439	ACN44439 standard; cDNA; 4684 BP.
XX	XX
XX	XX
XX	AC AC
XX	ACN44439;
XX	XX
DT	18-NOV-2004 (first entry)
XX	XX
DE	Human mRNA sequence hCT1952652.
XX	XX
KW	Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.
XX	XX
OS	Homo sapiens.
XX	XX
PN	WO2003073826-A2.
XX	XX
PD	12-SEP-2003.
XX	XX
PF	28-FEB-2003; 2003WO-US006235.
XX	XX
PR	01-MAR-2002; 2002US-00087192.
XX	XX
PA	(SAGR-) SAGRES DISCOVERY.
XX	XX
PI	Morris DW;
XX	XX
DR	WPI; 2003-328604/31.
XX	XX

```
QY 721 TCCAAATACAGTCTCTTAATGTTGAACA CTTGCACATCCGGAGTTCCACCGGTCAAAG 780
Db 2870 TCCAAATACAGTCTCTTAATGTTGAACA CTTGCACATCCGGAGTTCCACCGGTCAAAG 2929
QY 781 TGGAGACATCGGCATCGGCATCGAGCAGCAGATCCAGCTCGAGCTTTCAGCTTGGTCA 840
Db 2930 TGGAGACATCGGCATCGGCATCGAGCAGCAGATCCAGCTCGAGCTTTCAGCTTGGTCA 2989
QY 841 CCAAGACGGGAGCGCTGTCCTAGACATGGAGTGCAGACTCGGGCATCGACCTGC 900
Db 2990 CCAAGACGGGAGCGCTGTCCTAGACATGGAGTGCAGACTCGGGCATCGACCTGC 3049
QY 901 AGTGACACTGGCCCCCTGATGGCAGCTTCGCCCTGGAGCTGGAGGTTCAAGCATGCCCAGC 960
Db 3050 AGTGACACTGGCCCCCTGATGGCAGCTTCGCCCTGGAGCTGGAGGTTCAAGCATGCCCAGC 3109
QY 961 TGGAGAACAGGCCCTAACCCCTGCCCTCCACCGCGGCTCCACACTGCCGGAAGAGCCT 1020
Db 3110 TGGAGAACAGGCCCTAACCCCTGCCCTCCACCGCGGCTCCACACTGCCGGAAGAGCCT 3168
QY 1021 TCTGCTCGGTGCAGATGCTCCCTGAAACACAGGCTCAGCGCTTCCAGGGATVYG 1080
Db 3169 TCTGCTCGGTGCAGATGCTCCCTG--AAACACAGGCTCAGCGCTTCCAGGG--ATC 3224
QY 1081 NCCAGCCCCCGGCTCARCAGNTGGGAAC CAGGGCTTCGNCAGCAGCAGGTTNGGGG 1140
Db 3225 TGCCAGCCCCCGGCTCAGCAGTGGGAC CAGGGCTTCGNCAGCAGCAGG--GTGGG 3277
QY 1141 CAAGCNAGATGCTCCAGGATTTCA CACTGAGCCCTGCCAGCCTGCTGGAADA 1200
Db 3278 GCAAGCAAGATGCTCCAGGATTTCA CACTGAGCCCTGCCAGCCTGCTGGAADA 3334
QY 1201 AAAACATNCCGCCACGTTGAAGACAG AAGAGGATGNCAGGAGTTNACCTVGGGAA 1260
Db 3335 AACACT--CCGCCAAGTGAAGACA----GAGGAGATGGCAGAGTTACCTCGGGAA 3388
QY 1261 ACAAAACAGGATTTTNTTTCGCCCTG CTTCCAGTNCAGTGTGCTGACCCCTTGG 1320
Db 3389 CAACAGGATCTTCTCTGCCCTGCTCCAG-----TCAGTTGGCTGACCCGCTTGG- 3440
QY 1321 ANTCACTGACCATTTTGTGGCAGANCAG GGGAGACAGCTTCAGCCCTGGGTGAGAAGG 1380
Db 3441 -ATCAGTGACCATTTTGTCTGGCAGA- CAGGGGAGACAGCTTCAGCCCTGGGTGAGAAGG 3498
QY 1381 GTGGCGAGCCCTTCGGCCCTCACCCTN CAGAGTGTCTGTGNAGAGTCAAGTGTGTA 1440
Db 3499 GTGGCGAGCCCTTCGGCCCTCACCCT- CCAGGCTGTCTGTG-AGAGTGTCAAGTGTGTA 3556
QY 1441 AGGNCACAAANCTCAGNNTCAGTGCAGA ACCAGGTCAGCAGGTATGCCCGCCCGNTA 1500
Db 3557 AGGNCACAAACTC---AGGTTCAAGTGCAGAAC CAGGT-CAGCAGGTATGCCCGCCCGTAG 3612
QY 1501 GGTAAANNNGGGGCCCTCTNAAACCCCTT GCGTNGGCCCTNACCTNGGCCAGCTCANC 1560
Db 3613 GTTAA-----GGGGGCCCTCTAAACCCCTT GCGTNGGCCCTC---CCTGGCCAGCTCA-CCC 3664
QY 1561 CTTTTGGGTGTAGGGGAAAGAAATCCCTG AACCCTGGGAGGCTWCCTGGGTAGAAATACAC 1620
Db 3665 CTTTTGGGTGTAGGGGAAAGAAATCCCTG AACCCTGGGAGGCT-CCCTGGGTAGAAATACAC 3723
QY 1621 CACACTTTTCAAGTTGTGCAACACAGTCTG AGTTGACCTCTGTGGTTCAGCCCAAGGACC 1680
Db 3724 CACACTTTTCAAGTTGTGCAACACAGTCTG AGTTGACCTCTGTGGTTCAGCCCAAGGACC 3783
QY 1681 AAAGAAAGTGTAGTCAAGTGGTCTTCAGT NCCAGACATGTCGCCCTTTCGCTGCTG 1740
Db 3784 AAAGAAAGTGTAGTCAAGTGGTCTTCAGT NCCAGACATGTCGCCCTTTCGCTGCTG 3842
QY 1741 GCTACCACTCTTCCCAAGCAGCAGGCGCCG AGCCCTTCAGGCCCAGCAGCTGCCCCAG 1800
Db 3843 GCTACCACTCTTCCCAAGCAGCAGGCGCCG AGCCCTTCAGGCCCAGCAGCTGCCCCAG 3902
```

```
QY 1801 ACTCGCTGGCACA CTGATTTCCTCTCTCTA AAGGTGAAGGTGATGACAGGATATGCCCTGA 1860
Db 3903 ACTCGCTGGCACA CTGATTTCCTCTCTCTA AAGGTGAAGGTGATGACAGGATATGCCCTGA 3962
QY 1861 CAGGAACAGTCTGTGGATGGACATGATCAGT GCCTNAAGGNAAGCAGCAGACAGAGAGCGY 1920
Db 3963 CAGGAACAGTCTGTGGATGGACATGATCAGT GCCTNAAGGNAAGCAGCAGACAGAGAGAGCGC 4020
QY 1921 TCCGGCGCCCGCAGNCCCACTNATCAGTGT NCCAGCGTGTGCTGCTTNGTTCGCCAGNAGACA 1980
Db 4021 TCCGGCGCCCGCAGNCCCACTNATCAGTGT ----CCAGCGTGTGCTGCTTCCC-----CAGA 4069
QY 1981 GCTNCAGNCATCANCACATGACACTNACCT CCGCTNGCCCTGCGCCCGCANGAGGCTACTG 2040
Db 4070 GCACAGCTCAGCATCACA CTGACACTCACC CTGCGCTGCCCCCTGG--CCAGAGGGTACTG 4127
QY 2041 CCGNACGGCAGCTTTGCACTCTGATGATGNA CCTCAAAGCAGCTTTTCATGGCTNGCTCTTNG 2100
Db 4128 CCG-ACGGCAGCTTTGCACTCTGATG-TC TGA TG-ACCTCAAAGCAGCTTTTCATGGCTGCCCCCTGCGCA 4184
QY 2101 GCAGGNCAGGNCAGGNCAGTGCACANCTGT AGNAGCATANGCANGCCAGGAGATGG 2160
Db 4185 GGGCAGGCGCAGGGCAGTGCACACTGT-----AGGAGCATAGCAGCCAGGAGATGG 4234
QY 2161 GGTGNAAGGGGANCACAGTCTTTGAGTGTG CCACTGTCANCATGATGTGCTCAAACTCTTN 2220
Db 4235 GGTGAA--GGGACACAGCTTTGAGCTGTCCA -CATGATGTGACT-CCTCAAACTCTTT- 4289
QY 2221 NCCAGNATTTCTTAAGATAGCANCCTTNG CCCCCTTNGCCCATGCCCCAGCTTAGCCTTTCT 2280
Db 4290 --CCAGATTTCTTAAGATAGCACCCTCC -TTCCCATTTGCCCCAGCTTAGCCTTTCT 4345
QY 2281 CCCAGGGGAGCTANTCTCAGGACTCAGCTAG CATTAAATCAGCTGTGNAATCGTCAGGGGG 2340
Db 4346 CCCAGGGGAGCTA-CTCAGGACTCAGCTAG CATTAAATCAGCTGTG-AATCGTCAGGGGG 4403
QY 2341 TGTCTGTAGCTCAACTCTCTGGGCGAGGGG ACGCCAGACTCCGCTGGGAGAAAGCTCAT 2400
Db 4404 TGTCTGTAGCTCAACTCTCTCTGGGCGAG GGAGCCGAGACTCCGCTGGGAGAAAGCTCAT 4463
QY 2401 TCCACATCTTTGCCAAGCAGCCTTTNGTCC AGCTGTGCCAATTTGAGTTCAGACTGTCTCCC 2460
Db 4464 TCCACATCTTTGCCAAGCAGCCTTT-GTCC AGCTGTGCCAATTTGAGTTCAGACTGTCTCCC 4522
QY 2461 GGGCAGAGAGCCCGGCCCGCCAGCACATAA GAACTGCAGCCTTGGTACTGCAGAGTCTG 2520
Db 4523 GGGCAGAGAGCCCGGCCCGCCAGCACATAA GAACTGCAGCCTTGGTACTGCAGAGTCTG 4582
QY 2521 GGTGTAGAGAACTCTTTGTAAGCAATAAAG TTTGGGGTGTGATGACAAATGTT 2572
Db 4583 GGTGTAGAGAACTCTTTGTAAGCAATAAAG TTTGGGGTGTGATGACAAATGTT 4634
```

RESULT 8

ADQ67410

ID ADQ67410 standard; cDNA; 3152 BP.

XX AC ADQ67410;

XX DT 07-OCT-2004 (first entry)

XX DE Novel human cDNA sequence #2383.

ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
cancer.

XX OS Homo sapiens.

XX PN EP1440981-A2.

XX FD 28-JUL-2004.

XX PF 21-JAN-2004; 2004EP-00001196.
XX PR 21-JAN-2003; 2003JP-00102206.
XX PR 09-MAY-2003; 2003JP-00131392.
XX PA (REAS-) RES ASSOC.BIOTECHNOLOGY.
XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX PI Yamamoto J, Isono Y, Nagai K, Irie R;
XX DR WPI: 2004-535376/52.
XX DR P-PSDB; ADQ67717.
XX PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
XX PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX PS Claim 1; SEQ ID NO 4571; 2449pp; English.
XX CC The invention relates to 2495 novel polynucleotides (I) and their encoded
XX CC polypeptides, sequences hybridizing to these nucleotides, sequences
XX CC encoding partial polypeptides and sequences having 70% or 90% identity to
XX CC the nucleotide and protein sequences. The nucleotides and polypeptides
XX CC are useful as diagnostic markers or therapeutic target for the diseases
XX CC or morbid states. They are also useful for treating osteoporosis,
XX CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
XX CC dementia and various cancers. This sequence corresponds to a nucleotide
XX CC sequence of the invention.
XX SQ Sequence 3152 BP; 704 A; 1000 C; 874 G; 574 T; 0 U; 0 Other;

Query Match 67.3%; Score 1771.6; DB 12; Length 3152;
Best Local Similarity 89.2%; Pred. No. 0;
Matches 2287; Conservative 5; Mismatches 186; Indels 87; Gaps 35;

QY 1 CCCCTCTCACAGCCAGCCGCGCATCAAGAGGGGCTGAGGAAGAGCCCATCACCGCGTGT 60
DB 675 CCCCTCTCACAGCCAGCCGCGCATCAAGAGGGGCTGAGGAAGAGCCCATCACCGCGTGT 734

QY 61 CTGCAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120
DB 735 CTGCAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 794

QY 121 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC 180
DB 795 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC 854

QY 181 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTCCGCCAAGGGCCCCAGGGCCCCGGCCAG 240
DB 855 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTCCGCCAAGGGCCCCAGGGCCCCGGCCAG 914

QY 241 CTGAGGAGACAACAGGCGAGCCCTTAAGCTCCAGCCCTCTCTCCACAGAGCCCCCAG 300
DB 915 CTGAGGAGACAACAGGCGAGCCCTTAAGCTCCAGCCCTCTCTCCACAGAGCCCCCAG 974

QY 301 AGCCAAACAAAGTCTCTCCCTTGACTTTGACGAAGGAGGAGTCTGGGATGTGGGAACCT 360
DB 975 AGCCAAACAAAGTCTCTCCCTTGACTTTGACGAAGGAGGAGTCTGGGATGTGGGAACCT 1034

QY 361 TACCTCTGTCTCTCTGGAGCCAGCCCTCTGCCAAGAACCCAGCTACACGAGCGGAAG 420
DB 1035 TACCTCTGTCTCTCTGGAGCCAGCCCTCTGCCAAGAACCCAGCTACACGAGCGGAAG 1094

QY 421 CAACTGTCGGAGAGGAGTCTGAGGAGTGGGAATAGAAATATTCTCAACAGCCGTGT 480
DB 1095 CAACTGTCGGAGAGGAGTCTGAGGAGTGGGAATAGAAATATTCTCAACAGCCGTGT 1154

QY 481 CCCAGCCATTTCTCTGGAGGAGCAGGAGAAATTTCTCTGTGCTCAGCATCGACAGCC 540
DB 1155 CCCAGCCATTTCTCTGGAGGAGCAGGAGAAATTTCTCTGTGCTCAGCATCGACAGCC 1214

QY 541 TCTCCCTGTCCGATGACAGTGAGGAAGAACCCATCAAAAGGCTCTCAAAAGCTTCGGGGACA 600
DB TCTCCCTGTCCGATGACAGTGAGGAAGAACCCATCAAAAGGCTCTCAAAAGCTTCGGGGACA 2308

DB 1215 TCTCCCTGTCCGATGACAGTGAGGAAGAACCCATCAAAAGGCTCTCAAAAGCTTCGGGGACA 1274
QY 601 CCCTGAGCTCAGGCGTACACTCTCTGAGCAGCAGGCGCGAGGCTCGAAGCTCCAGCTCGA 660
DB 1275 CCCTGAGCTCAGGCGTACACTCTCTGAGCAGCAGGCGCGAGGCTCGAAGCTCCAGCTCGA 1334
QY 661 ACATGTGTGCGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTGTGTGAAG 720
DB 1335 ACATGTGTGCGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTGTGTGAAG 1394
QY 721 TCCAAATACAGTCTCTTAATGTGTGAACACCTGACATCCCGGAGTTCCACCGGTCAAG 780
DB 1395 TCCAAATACAGTCTCTTAATGTGTGAACACCTGACATCCCGGAGTTCCACCGGTCAAG 1454
QY 781 TGGGAGACATCGCCACTTGGCATCAGCAGCAGATCCAGCTTCAGCTTCAGCTTGGTCA 840
DB 1455 TGGGAGACATCGCCACTTGGCATCAGCAGCAGATCCAGCTTCAGCTTCAGCTTGGTCA 1514
QY 841 CCAAAGACGGGCGAGCTTTCGCTACGATGAGAGTGCAGACTCGGGCATCGACCTGC 900
DB 1515 CCAAAGACGGGCGAGCTTTCGCTACGATGAGAGTGCAGACTCGGGCATCGACCTGC 1574
QY 901 AGTGCACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGAGGCTCAAGCATGGCCAGC 960
DB 1575 AGTGCACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGAGGCTCAAGCATGGCCAGC 1634
QY 961 TGCAGAACAGCGCCCTAACCCCTGCCCTCCACCGCGGCTCCACACTGCGG-AAAGAGCT 1020
DB 1635 TGCAGAACAGCGCCCTAACCCCTGCCCTCCACCGCGGCTCCACACTGCGG-AAAGAGCT 1693
QY 1021 TCTGTCTCGGTGACGATGCTGCCCTGAAAAACAGCTCAGCCGTTCCAGGGGATYTG 1080
DB 1694 TCTGTCTCGGTGACGATGCTGCCCTG-AAAACAGAGCTCAGCCGTTCCAGGG-ATC 1749
QY 1081 NCCAGCCCCCGGCTCAGCAGNTGGGAAACAGGGCTCGNACGACGAGTNGGGG 1140
DB 1750 TGCAGCCCCCGGCTCAGCAGTGGGAAACAGGGCTCGCAGCAGCAG-ATG-ATG 1802
QY 1141 CAAGCAGAAATGCTCCAGGATTTACANCTGAGCCNTGCCANCTCTGCTGAADA 1200
DB 1803 GCAAGCAGATGCTCCAGGATTTACACC--TGAGCCCTGCCACCTGCTGAGAA 1859
QY 1201 AAACATNCCCGCAGCTGAAGACAGAGAGGATCGNACGAGTNNACTYVGGGAA 1260
DB 1860 AACACT--CGCCACCTGAAGAGACA---GAGGAGGATGCGAGAGTTACCTCGGAAA 1913
QY 1261 ACAAACAGGATCTTNTTCTGCCCTGCTCCAGTNCAGTGGCCCTGACCCGCTTG 1320
DB 1914 CAAACAGGATCTTCTGCCCCCTGCTCAG-----TCGAGTTGGCCTGACCCGCTT--G 1964
QY 1321 ANTCACTGACCATTTGTTGGCAGANCAAGGGGAGAGCAGCTTCAGCCCTGGGTCAAGAGG 1380
DB 1965 GATCAGTGACCATTTGCTGGCAGA-CHAGGGAGAGCAGCTTCAGCTGGGTCAAGAGG 2023
QY 1381 GTGGGCGAGCCCTTCGGCCCTCACCCTNCCAGGCTGCTGTGNAGAGTGTCAAGTGTGA 1440
DB 2024 GTGGGCGAGCCCTTCGGCCCTCACCCT-CCAGGCTGCTGTG-AGAGTGTCAAGTGTGA 2081
QY 1441 AGGNCCTCAANCTCAGGNTTCAAGTGAAGAACAGGTCNACAGGATGTCGCCGCCGNTA 1500
DB 2082 AGGGCCCCAACTC---AGGTTCAAGTGAAGAACAGGTCNACAGGATGTCGCCGCCGNTA 2137
QY 1501 GGTAAANNNGGGGCCCCCTTNAACCCCTTCCCTNCGGCTNCACTNCGGCGCAGCTCANCCC 1560
DB 2138 GTTAA---GGGGGCCCCCTTAAACCCCTTGCCTGGCCCTCA---CCTGCCAGCTCA-CCC 2189
QY 1561 CTTTGTGGGTAGGGGAAAAGAAATGCTGACCCCTGGGAAGGCTTCCCTGGTGTAGTAATAC 1620
DB 2190 CTTTGTGGGTAGGGGAAAAGAAATGCTGACCCCTGGGAAGGCT-CCCTGGTGTAGTAATAC 2248
QY 1621 CACATTTTTCAGGTTGTGCAACACAGGTCCTGAGTTGACCTCTGTTTTCAGCCAGGACC 1680
DB 2249 CACATTTTTCAGGTTGTGCAACACAGGTCCTGAGTTGACCTCTGTTTTCAGCCAGGACC 2308

1681 AAAGAAGGTGTGTAAGTGAAGTGTCTTCAGTNCCTCCAGACATGTGCCCTTTGCTGCTG 1740
2309 AAAGAAGGTGTGTAAGTGAAGTGTCTTCAGT-CCCAGACATGTGCCCTTTGCTGCTG 2367
1741 GCTACCACTTTCTCCAGAGCAGCAGGCCCGAGGCCCTTTCAGGCCCCAGCACTGCCCCAG 1800
2368 GCTACCACTTTCTCCAGAGCAGCAGGCCCGAGGCCCTTTCAGGCCCCAGCACTGCCCCAG 2427
1801 ACTCGTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGGTGATGCGAGGATATGCTCTGA 1860
2428 ACTCGTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGGTGATGCGAGGATATGCTCTGA 2487
1861 CAGGAACAGTCTGTGATGGACATGATCAGTCTTNAAGNAAGCAGCAGAGAGAGCY 1920
2488 CAGGAACAGTCTGTGATGGACATGATCAGTCTTAA--GGAAGCAGCAGAGAGAGCG 2545
1921 TCCGGGCCCCAGNCCCCACTNATCAGTGTNCCAGCGTGTCTNGTGTNCCCCAGNAGCACA 1980
2546 TCCGGGCCCCAGCCTATCAGTGT--CCAGCGTGTGTTCCC-----CAGA 2594
1981 GCTNCAGNATCAGCACTGACACTNACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
2595 GCACAGCTCAGCATCAGTACACTCAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2652
2041 CCGNACGGCACTTTGACNCTCTGATGNACCTCAAAGCACTTTTCATGGCTNGCCTCTNNG 2100
2653 CCG-ACGGCACTTTGAC-TCGTATG-ACCTCAAAGCACTTTTCATGGCTGCTGCTGCTG 2709
2101 CGAGGNCAGGNCAGGNCAGTACACTGTAGNAGCATANGAANGCCAGGAGATGG 2160
2710 GCGCAGGCGAGGCGAGTACACTGT-----AGGAGCATACCAAGCCAGGAGATGG 2759
2161 GGTGNAAGGNCAGTCTTGAAGTGTCCANCATGTGATGACTNCTCAAACTCTTN 2220
2760 GGTGAAGGG--ACACAGTCTTGAAGTGTCCA-CATGCTGTGACT-CCTCAAACTCTT- 2814
2221 NCCAGNATTTCTTAAGATACANCCCTTNCCTTCCCATTTGCCAGCTTACGCTCTTCT 2280
2815 --CCAGATTTCTTAAGATACACACCC--TTCCCATTTGCCAGCTTACGCTCTTCT 2870
2281 CCCAGGGAGCTANTCAGGACTCAGTAGCAATTAATCAGCTGTGNAATCGTCAGGGGG 2340
2871 CCCAGGGAGCTA-CTCAGGACTCAGTAGCAATTAATCAGCTGTG-AATCGTCAGGGGG 2928
2341 TGCTCTGCTAGCTCAACCTCTTGGGCGAGGGGACCCGAGACTCGTGGGAGAGCTCAT 2400
2929 TGCTCTGCTAGCTCAACCTCTTGGGCGAGGGGACCCGAGACTCGTGGGAGAGCTCAT 2988
2401 TCCACATCTTGGCCAGACAGCTTTNGTCCAGCTGTCCACATTCAGTTCAGACTGCTCCC 2460
2989 TCCACATCTTGGCCAGACAGCTTTT-GTCCAGCTGTCCACATTCAGTTCAGACTGCTCCC 3047
2461 GGGGAGAGAGCCCCGGCCCCAGCACAATAAGAACTGCAGCTTGGTACTGCAGAGTCTG 2520
3048 AGGGAGAGAGCCCCGGCCCCAGCACAATAAGAACTGCAGCTTGGTACTGCAGAGTCTG 3107
2521 GGTGTAGAGACTCTTTGTAGCAATAAAGTTTGGGGTGTATGAC 2565
3108 GGTGTAGAGACTCTTTGTAGCAATAAAGTTTGGGGTGTATGAC 3152

RESULT 9
AAV73917
ID AAV73917 standard; cDNA; 3156 BP.
XX
AC AAV73917;
XX
DT 04-MAR-1999 (first entry)
XX Human NIK cDNA.
XX
KW NIK; Nuclear factor-kappa B; NF-kB; signal transduction; TNF; human;

KW tumour necrosis factor; NF-kB-inducing kinase; screening; interaction;
KW modulator; immune response; inflammatory response; viral gene;
XX diagnostic; therapy; ds.
OS Homo sapiens.
XX
XX US5854003-A.
XX
XX 29-DEC-1998.
XX
XX 26-FEB-1998; 98US-00032475.
XX
XX 03-JUL-1997; 97US-00887518.
XX (TULA-) TULARIK INC.
XX
XX Wu L, Rothe M;
XX WPI; 1999-094902/08.
XX
XX Screening agents for modulating interaction of nuclear factor kappaB
XX inducing kinase - with kinase-binding target, useful for controlling
XX levels of the kinase, for treatment and diagnosis of conditions
XX associated with e.g. inhibition of signal transduction by tumour necrosis
XX factor.
XX
XX Disclosure; Col 11-14; 16pp; English.
XX
XX This sequence encodes a novel human nuclear factor-kappaB (NFkB)-inducing
XX kinase (NIK) which is used in a method for screening for agents that
XX modulate the interaction of NIK with a NIK-binding target. The encoded
XX protein can be used as a modulator of cellular functions at the NIK
XX level, or for development of such compounds. NFkB is involved in
XX expression of many immune and inflammatory responses and of some
XX important viral genes. The protein may be used diagnostically and
XX therapeutically, in conditions associated with abnormal utilisation of
XX pathways that involve NFkB, e.g. inhibition of signal transduction by
XX tumour necrosis factor (TNF)
XX
XX Sequence 3156 BP; 736 A; 991 C; 895 G; 534 T; 0 U; 0 Other;

Query Match 43.1%; Score 1133.4; DB 2; Length 3156;
Best Local Similarity 94.8%; Pred. No. 2.1e-305;
Matches 1241; Conservative 4; Mismatches 43; Indels 21; Gaps 7;
Qy 1 CCCCTCTCACAGCCCCAGGCCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
Db 1868 CCCCTCTCACAGCCCCAGGCCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 1927
Qy 61 CTGCAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120
Db 1928 CTGCAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 1987
Qy 121 GCCCTTGGAGGGGAGAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCCAATTACC 180
Db 1988 GCCCTTGGAGGGGAGAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCCAATTACC 2047
Qy 181 ACAGACCTTCATGCCAGCCGAGAGAGCTTTCCCAAGGGCCCCAGGGCCCCCGGCAG 240
Db 2048 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTCCCAAGGGCCCCAGGGCCCCCGGCAG 2107
Qy 241 CTGAGGAGACAAACAGGCGAGAGCCCTTAAGTCTCCAGCTCTCTCCACAGAGCCCCCAG 300
Db 2108 CTGAGGAGACAAACAGGCGAGAGCCCTTAAGTCTCCAGCTCTCTCCACAGAGCCCCCAG 2167
Qy 301 AGCCAAAACAGTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCCCT 360
Db 2168 AGCCAAAACAGTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCCCT 2227
Qy 361 TACCTCTGCTCTCTCTGGAGCGAGCCCTTGCAGAAACCCAGCTTCCAGAGCGGAAAG 420
Db 2228 TACCTCTGCTCTCTCTGGAGCGAGCCCTTGCAGAAACCCAGCTTCCAGAGCGGAAAG 2287

```
Qy 421 CAACGCTCCCGAGGAGGAACTGAGCAGCTGGAAATAGAAATATTCTCTCAACAGCCTGT 480
Db 2288 CAACGCTCCCGAGGAGGAACTGAGCAGCTGGAAATAGAAATATTCTCTCAACAGCCTGT 2347
Qy 481 CCCAGCATTCTCTGAGGAGCAGGAGCAAAATCTCTGCTGCTCAGAGCTCGACAGCC 540
Db 2348 CCCAGCATTCTCTGAGGAGCAGGAGCAAAATCTCTGCTGCTCAGAGCTCGACAGCC 2407
Qy 541 TCTCCCTGTGCGATGACAGTGAGAGAAACCCATCAAAAGGCTCTCAAAAGCTCGCGGACA 600
Db 2408 TCTCCCTGTGCGATGACAGTGAGAGAAACCCATCAAAAGGCTCTCAAAAGCTCGCGGACA 2467
Qy 601 CCCTGAGCTCAGGCGTATACCTCTGAGAGCAGCCAGGCGGAGCTCGAAGCTCCAGCTGGA 660
Db 2468 CCCTGAGCTCAGGCGTATACCTCTGAGAGCAGCCAGGCGGAGCTCGAAGCTCCAGCTGGA 2527
Qy 661 ACATGGTGTGCGCGCGGGCGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 720
Db 2528 ACATGGTGTGCGCGCGGGCGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 2587
Qy 721 TCCAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCACCGGCTCAAAG 780
Db 2588 TCCAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCACCGGCTCAAAG 2647
Qy 781 TGGAGACATCGCACCTGGCATACAGCAGCCAGATCCAGCTGAGCCTTCAGCTTGTGCA 840
Db 2648 TGGAGACATCGCACCTGGCATACAGCAGCCAGATCCAGCTGAGCCTTCAGCTTGTGCA 2707
Qy 841 CCAAGACGGGCGAGCTGTTGCTACGACATGGAGTTCAGACTCGGGATCGACCTGC 900
Db 2708 CCAAGACGGGCGAGCTGTTGCTACGACATGGAGTTCAGACTCGGGATCGACCTGC 2767
Qy 901 AGTGACACTGGCCCTGATGGGAGCTTGGCTGGAGCTGAGGCTCAAGCATGGCCAGC 960
Db 2768 AGTGACACTGGCCCTGATGGGAGCTTGGCTGGAGCTGAGGCTCAAGCATGGCCAGC 2827
Qy 961 TGGAGAACAGGCCCTAACCTCCCTCCACCGCGGCTCCACACTGCCGGAAGACGCCT 1020
Db 2828 TGGAGAACAGGCCCTAACCTCCCTCCACCGCGGCTCCACACTGCCGGAAGACGCCT 2887
Qy 1021 TCCTGCTCGTGCAGATGCTGCCCTGAAACACAGAGCTCAGCGTTCCAGGGGATYTG 1080
Db 2888 TCCTGCTCGTGCAGATGCTGCCCTGAAACACAGAGCTCAGCGTTCCAGGGGATYTG 2945
Qy 1081 NCCAGCCCCCGGCTCARGCTGGGAACACAGGCGCTCGNACGACGNAAGTNGGGG 1140
Db 2946 GCCAGCCCCCGGCTCA--CAGTGGAAACAGGCGCTCG-----CAGCAGAGGTGGG 2998
Qy 1141 CAAGCAGAAATGCTCCCGAGATTTCACAGCTGAGCCGCTGAGCCGCTGCTGAADA 1200
Db 2999 GCAAGCAGAAATGCTCCCGAGATTTCACAGCTGAGCCGCTGAGCCGCTGCTGAADA 3052
Qy 1201 AAACAYTNCGCCAGCTGAGAGACAGAGAGATGAGGATGAGGTTNACCTYGGGAA 1260
Db 3053 AAACAYTNCGCCAGCTGAGAGACAGAGAGATGAGGATGAGGTTNACCTYGGGAA 3108
Qy 1261 ACAAACAGGATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1309
Db 3109 ACAAACAGGATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3155
```

RESULT 10

AAV71603

ID AAV71603 standard; cDNA; 3156 BP.

XX AAV71603;

AC AAV71603;

XX 04-FEB-1999 (first entry)

DE Human NF-kB-inducing kinase (NIK) polypeptide encoding cDNA.

XX NIK, IkB; TNF; kinase; nuclear factor kappa B; inhibition;
KW tumour necrosis factor; binding; genetic hybridisation; screening;
KW

signal transduction; biopharmaceutical; human; de.

Homo sapiens.

Key Location/Qualifiers

CDS 1..2844

/tag= a

/product= "NIK polypeptide"

XX US5843721-A.

XX 01-DEC-1998.

XX 03-JUL-1997; 97US-00887518.

XX 03-JUL-1997; 97US-00887518.

XX (TULA-) TULARIK INC.

XX Wu L, Rothe M;

XX WPI; 1999-044580/04.

XX P-PSDB; AAW81564.

XX Probe, vector or recombinant nucleic acid encoding a polypeptide, especially human nuclear factor kappa-B-inducing kinase protein - useful for producing recombinant protein.

XX Claim 1; Col 11-14; 15pp; English.

XX This cDNA encodes a nuclear factor kappa B (NF-kB)-inducing kinase (NIK) polypeptide. The NIK polypeptide or its fragment has one or more activities selected from kinase activity and inhibitory activity; IkB kinase-alpha and beta binding activity and binding inhibitory activity; tumour necrosis factor (TNF) receptor-associated factor 2 binding activity and binding inhibitory activity; IkB binding activity and binding inhibitory activity; NF-kB activating and inhibitory activity. A vector containing the NIK nucleic acid can be used to transform host cells for the recombinant production of the protein. The NIK nucleic acid and the polypeptide may be used in diagnosis (e.g. genetic hybridisation screen for NIK transcripts), therapy (e.g. NIK kinase inhibitors to inhibit TNF signal transduction), and in the biopharmaceutical industry

XX Sequence 3156 BP; 736 A; 991 C; 895 G; 534 T; 0 U; 0 Other;

Query Match 43.1%; Score 1133.4; DB 2; Length: 3156;

Best Local Similarity 94.8%; Pred. No. 2.1e-305;

Matches 1241; Conservative 4; Mismatches 43; Indels 21; Gaps 7;

Qy 1 CCCTCTCACAGCCCGGCGCATCCAGAGGGGCTGAGGAAAGAGCCCATCCAGCGCTGT 60

Db 1868 CCCTCTCACAGCCCGGCGCATCCAGAGGGGCTGAGGAAAGAGCCCATCCAGCGCTGT 1927

Qy 61 CTCAGCGGAGCTGGAGGAAAGTGAACCGGGCACTACAGCAAGTGGAGGCTGAAGA 120

Db 1928 CTCAGCGGAGCTGGAGGAAAGTGAACCGGGCACTACAGCAAGTGGAGGCTGAAGA 1987

Qy 121 GCCCTTGGAGGGAGATATAAGAACCAAGACATCCACGCCCAATCAAGCCATATACC 180

Db 1988 GCCCTTGGAGGGAGATATAAGAACCAAGACATCCACGCCCAATCAAGCCATATACC 2047

Qy 181 ACCAGACCTCTCCATGCCCGGAGAGAGCTTTTCGCCAAGGGCCCAAGGGCCCGGCGAG 240

Db 2048 ACCAGACCTCTCCATGCCCGGAGAGAGCTTTTCGCCAAGGGCCCAAGGGCCCGGCGAG 2107

Qy 241 CTGAGGAGCAACAGGCGAGAGCCCTTAAGCTCCAGCTCTCTCCACACAGAGCCCGCAG 300

Db 2108 CTGAGGAGCAACAGGCGAGAGCCCTTAAGCTCCAGCTCTCTCCACACAGAGCCCGCAG 2167

Qy 301 AGCCAAACAGTCTCTCTCTGACTTTGAGCAAGGAGGAGTCTGGATGTGGAAACCT 360

Db 2168 AGCCAAACAGTCTCTCTCTGACTTTGAGCAAGGAGGAGTCTGGATGTGGAAACCT 2227

```
QY 361 TACCTCTGCTCCCTCGAGCAGAGCCCTCGCAGAAACCCCGAGCTCACACAGCGGAAAG 420
Db 2228 TACCTCTGCTCCCTCGAGCAGAGCCCTCGCAGAAACCCCGAGCTCACACAGCGGAAAG 2287
QY 421 CAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAAATATTCTTCAACAGCCTGT 480
Db 2288 CAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAAATATTCTTCAACAGCCTGT 2347
QY 481 CCAGCATTATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTCCTCAGCATCGACAGCC 540
Db 2348 CCAGCATTATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTCCTCAGCATCGACAGCC 2407
QY 541 TCTCCCTGTCGATGACATGATGAGAGAAACCCATCAAAAGCCTCTCAAAAGCTCGCGGACA 600
Db 2408 TCTCCCTGTCGATGACATGATGAGAGAAACCCATCAAAAGCCTCTCAAAAGCTCGCGGACA 2467
QY 601 CCTGAGCTCAGCGGTACACTCTCTGGAGCAGCAGCCGAGGCTCGAAGCTCCAGCTGGA 660
Db 2468 CCTGAGCTCAGCGGTACACTCTCTGGAGCAGCAGCCGAGGCTCGAAGCTCCAGCTGGA 2527
QY 661 ACATGCTGTGCGCGCGGCGGCCACCGACACCCCAAGCTATTTCATGGTGTGAAAG 720
Db 2528 ACATGCTGTGCGCGCGGCGGCCACCGACACCCCAAGCTATTTCATGGTGTGAAAG 2587
QY 721 TCCAAATACAGTCTCTTAATGTGAAACACCTGCAATCCGGGAGTTCCACCGGGTCAAAG 780
Db 2588 TCCAAATACAGTCTCTTAATGTGAAACACCTGCAATCCGGGAGTTCCACCGGGTCAAAG 2647
QY 781 TGGGAGACATCGCCACTGTCATCAGCAGCCAGATCCAGCTCGACCTTCAGCTTGGTCA 840
Db 2648 TGGGAGACATCGCCACTGTCATCAGCAGCCAGATCCAGCTCGACCTTCAGCTTGGTCA 2707
QY 841 CCAAGAGCGGCGAGCTGTTCTGCTACGATGAGGTGCCAGCTCGGGCATCGACCTGC 900
Db 2708 CCAAGAGCGGCGAGCTGTTCTGCTACGATGAGGTGCCAGCTCGGGCATCGACCTGC 2767
QY 901 AOTGCACATGCGCCCTGATGACAGCTGCTGCTGAGCTGAGGTCAAGCATGCCAGC 960
Db 2768 AGTGACACATGCGCCCTGATGACAGCTGCTGCTGAGCTGAGGTCAAGCATGCCAGC 2827
QY 961 TGGAGAACAGGCGCTTAACTGCTCCCTCCACCGCGGCTCCACCTGCGGAAAGCAGCT 1020
Db 2828 TGGAGAACAGGCGCTTAACTGCTCCCTCCACCGCGGCTCCACCTGCGGAAAGCAGCT 2887
QY 1021 TCTGCTCGGTGACAGTGTCTGCTGAAACACAGGCTCAGCCGTTCCAGGGGATGTG 1080
Db 2888 TCTGCTCGGTGACAGTGTCTGCTGAAACACAGGCTCAGCCGTTCCAGGGGATGTG 2945
QY 1081 NCCAGCCCCCGGCTCAGCAGNTGGAACCCAGGGCTTCGNCAGCNAGCNAAAGTNGGGG 1140
Db 2946 GCCAGCCCCCGGCTCA--CAGTGGGAACCCAGGGCTCG-----CAGCAGCAAGGTGGG 2998
QY 1141 CAAAGCAGAGTCCCTCCAGGATTTTCAACCTGAGCCGCTGCCCCCAGCTGCTGAADA 1200
Db 2999 GCAAGCAGAGTCCCTCCAGGATTTTCAACCTGAGCCGCTGCCCCCAGCTGCTGCTGA 3052
QY 1201 AAACATTCGCGCAGTGAAGAGACAGAGGAGTGCNAGGAGTTTNNACCTGCGGAA 1260
Db 3053 AAACATTCGCGCAGTGAAGAGACAGAGGAGTGCNAGGAGTTTNNACCTGCGGAA 3108
QY 1261 ACAAACAGGATCTTTTTCGCCCCCTGCTCCAGTNCAGTTCGCTG 1309
Db 3109 ACAAACAGGATCTTTTTCGCCCCCTGCTCCAGTNCAGTTCGCTG 3155
```

RESULT 11

AAV69285

ID AAV69285 standard; cDNA; 3156 BP.

XX

AC AAV69285;

XX

DT 04-FEB-1999 (first entry)

XX

DE Human NF-kB-inducing kinase (NIK) polypeptide encoding cDNA.

XX NIK; Ikb; NF-kB; TNF; kinase; nuclear factor kappa B; inhibition;
KW tumour necrosis factor; binding; genetic hybridisation; screening;
KW signal transduction; biopharmaceutical; immunogen; pharmacological;
KW transcription regulator; human; ds.

OS Homo sapiens.

XX Key Location/Qualifiers
FH 1..2844

FT /*tag= a
FT /product= "NIK polypeptide"

XX US5844073-A.

XX 01-DEC-1998.

XX 13-FEB-1998; 98US-00023321.

XX 03-JUL-1997; 97US-00887518.

XX (TULA-) TULARIK INC.

XX Wu L, Rothe M;

XX WPI; 1999-044664/04.

XX P-PSDB; AAW81561.

XX New isolated peptide comprising a specified 947 amino acid sequence - has
PT e.g. kinase activity, kinase inhibitory activity, Ikb kinase-alpha
PT binding activity, and Ikb kinase-alpha binding inhibitory activity.

XX Disclosure; Col 11-14; 15pp; English.

XX This cDNA encodes a nuclear factor kappa B (NF-kB)-inducing kinase (NIK)
CC polypeptide. The NIK polypeptide or its fragment has one or more
CC activities selected from kinase activity and inhibitory activity; Ikb
CC kinase-alpha and beta binding activity and binding inhibitory activity;
CC tumour necrosis factor (TNF) receptor-associated factor 2 binding
CC activity and binding inhibitory activity; Ikb binding activity and
CC binding inhibitory activity, NF-kB activating and inhibitory activity.
CC The NIK nucleic acid and the polypeptide may be used in diagnosis (e.g.
CC genetic hybridisation screen for NIK transcripts), therapy (e.g. NIK
CC kinase inhibitors to inhibit TNF signal transduction), and in the
CC biopharmaceutical industry (e.g. as immunogens, reagents for isolating
CC other transcription regulators, and reagents for screening chemical
CC libraries for pharmacological agents)

XX Sequence 3156 BP; 736 A; 991 C; 895 G; 534 T; 0 U; 0 Other;

Query Match 43.1%; Score 1133.4; DB 2; Length 3156;
Best Local Similarity 94.8%; Pred. No. 2.le-305;
Matches 1241; Conservative 4; Mismatches 43; Indels 21; Gaps 7;

QY 1 CCCCTCTCACAGCCAGCGCCATCCAGAGGGCTGAGGAAAGAGCCCATCCACGCGCTGT 60

Db 1868 CCCCTCTCACAGCCAGCGCCATCCAGAGGGCTGAGGAAAGAGCCCATCCACGCGCTGT 1927

QY 61 CTGCAGCGGAGCTGGGAGGAAAGGTGAACCCGGGCACCTACAGCAAGTGGGAGGTCTGAAGA 120

Db 1928 CTGCAGCGGAGCTGGGAGGAAAGGTGAACCCGGGCACCTACAGCAAGTGGGAGGTCTGAAGA 1987

QY 121 GCCCTTGGAGGGGAGATATAAGAACCAAGACATCCAGCCCAATCAAGCCATATACC 180

Db 1988 GCCCTTGGAGGGGAGATATAAGAACCAAGACATCCAGCCCAATCAAGCCATATACC 2047

QY 181 ACCAGACCTTCCATGCCAGCCAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCGGCCAG 240

Db 2048 ACCAGACCTTCCATGCCAGCCAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCGGCCAG 2107

QY 241 CTGAGGAGACAACAGGCGAGGCCCTTAAGTCTCAGGCTCTCTCCACACAGAGCCCCCAG 300

Db 2108 CTGAGGACAAACAGGACAGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCCCAG 2167
Qy 301 AGCCAAACAAAGTCTCTCCCTTGATCTTTGACAAAGGAGGAGTCTGGATGTGGAAACCT 360
Db 2168 AGCCAAACAAAGTCTCTCCCTTGATCTTTGACAAAGGAGGAGTCTGGATGTGGAAACCT 2227
Qy 361 TACCTCTGTCTCTCTCTGGAGCCAGCCCTTGCAGAAACCCAGCTTACAGAGCGGAAAG 420
Db 2228 TACCTCTGTCTCTCTCTGGAGCCAGCCCTTGCAGAAACCCAGCTTACAGAGCGGAAAG 2287
Qy 421 CAACCGTCCCGAGCAGGAACCTGACAGAGCTGGAAATAGAAATTTCTCAACAGCCTGT 480
Db 2288 CAACCGTCCCGAGCAGGAACCTGACAGAGCTGGAAATAGAAATTTCTCAACAGCCTGT 2347
Qy 481 CCCAGCCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTGCTCTCAGCATCGACAGCC 540
Db 2348 CCCAGCCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTGCTCTCAGCATCGACAGCC 2407
Qy 541 TCTCTCTGTCTGGATGACAGTGAGAGAACCCATCAAAAGGCTCTCAAAAGCTTCGGGGACA 600
Db 2408 TCTCTCTGTCTGGATGACAGTGAGAGAACCCATCAAAAGGCTCTCAAAAGCTTCGGGGACA 2467
Qy 601 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTTCAAGCTTGA 660
Db 2468 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTTCAAGCTTGA 2527
Qy 661 ACATGGTGTCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTCAATGGTGTGAAAG 720
Db 2528 ACATGGTGTCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTCAATGGTGTGAAAG 2587
Qy 721 TCCAAATACAGTCTTTAATGTTGAACACCTGACATCCGGGAGTTCACCGGGTCAAG 780
Db 2588 TCCAAATACAGTCTTTAATGTTGAACACCTGACATCCGGGAGTTCACCGGGTCAAG 2647
Qy 781 TGGGAGACATCGCCACTGCGATCAGCAGCCAGATCCCGAGCTCGAGCCTTCAGCTTGTGTA 840
Db 2648 TGGGAGACATCGCCACTGCGATCAGCAGCCAGATCCCGAGCTCGAGCCTTCAGCTTGTGTA 2707
Qy 841 CCAAAGCGGCGACCTGTTGCTTACACATGAGAGTGGCAGAGTCCGGGATCGACCTGC 900
Db 2708 CCAAAGCGGCGACCTGTTGCTTACACATGAGAGTGGCAGAGTCCGGGATCGACCTGC 2767
Qy 901 AGTGCACTGGCCCTGTATGGAGCTTGGCTGGAGCTGGAGGTTCAAGCATGGCCAGC 960
Db 2768 AGTGCACTGGCCCTGTATGGAGCTTGGCTGGAGCTGGAGGTTCAAGCATGGCCAGC 2827
Qy 961 TGGAGAACGCGCTTACCTCCCTCCAGCGCGGCTCCACACTGCGGGAAGAGCAGCT 1020
Db 2828 TGGAGAACGCGCTTACCTCCCTCCAGCGCGGCTCCACACTGCGGGAAGAGCAGCT 2887
Qy 1021 TCCTGCTCGGTGACAGATGCTGCCCTGAAACACAGGCTCAGCGCTTCCAGGGGATYTG 1080
Db 2888 TCCTGCTCGGTGACAGATGCTGCCCTGAAACACAGGCTCAGCGCTTCCAGGGGATYT-- 2945
Qy 1081 NCAGACCCCGGCTCAGCAGTGGGAAACAGGCGCTCGCAGCAGCAGTTCAGCGCTTCCAGGGGAT 1140
Db 2946 GCCAGCCCGGCTCA--CAGTGGGAAACAGGCGCTCG----CAGCAGCAAGTGGGG 2998
Qy 1141 CAACGACAGTCTCCAGGATTTACACCTGAGCCCTGAGCCCTGCCCCAGCCCTGCTGADA 1200
Db 2999 GCAAGCAGATGCTCCAGGATTTACACTGAGCCCTGCCCCAC-----CCTGCTGA 3052
Qy 1201 AAACATTCGCCACGCTGAAGAGACAGAGAGTGGAGGATGNNACCTTGGGAA 1260
Db 3053 AAACATTCGCCACGCTGAAGAGACAGAGAGTGGAGGATGNNACCTTGGGAA 3108
Qy 1261 AAAAAACAGGGAATTTTCTGCCCCCTGCTCCAGTTCAGTTCGAGTTCGCTG 1309
Db 3109 AAAAAACAGGGAATTTTCTGCCCCCTGCTCCAGTTCAGTTCGAGTTCGCTG 3155

ID XX AAX87842 standard; DNA; 2844 BP.
AC AAX87842;
DT XX 09-NOV-1999 (first entry)
DE XX NF-kB inducing kinase (NIK) DNA.
XX NIK; NF-kB inducing kinase; inflammation; tumour necrosis factor;
KW interleukin-1; cytokine; inhibitor; antiinflammatory; apoptosis;
KW anti-apoptotic; human; ss.
OS Homo sapiens.
XX WO9943704-A1.
PN 02-SEP-1999.
XX 25-FEB-1999; 99WO-US004110.
XX 27-FEB-1998; 98US-0076299P.
XX (REGC) UNIV CALIFORNIA.
PI Greene WC, Lin X, Gelezuinas R;
XX WPI; 1999-518837/43.
DR P-PSDB; AAY31665.
XX New NF-kB inducing kinase proteins, used for inhibiting a kB-dependent
PT immune response, e.g. an inflammatory response or an anti-apoptotic
PT response.
XX Disclosure; Fig 10; 48pp; English.
XX This DNA sequence codes for human wild-type NIK (see AAY31665), a NF-kB
CC inducing kinase having serine/threonine kinase activity. The invention
CC relates to: (a) an N-terminus deletion mutant NIK protein (see AAY31666);
CC and (b) a kinase deficient NIK mutant protein (see AAY31667 and AAY31668)
CC that inhibits auto-phosphorylation or transphosphorylation. The invention
CC provides the molecular basis for cytokine induction of NF-kB-dependent
CC immune and inflammatory responses, emphasising a role for both NIK-NIK
CC and NIK-IKK (IKK-specific kinase) interactions. A novel and highly
CC specific method for modulating NF-kB-dependent immune, inflammatory and
CC anti-apoptotic responses is based on interruption of the critical protein
CC interaction of NIK and IKK. The mutant NIK proteins are used in claimed
CC methods for inhibiting NF-kB-dependent gene expression. The kinase-
CC deficient NIK mutant proteins inhibit activation of IKK. The N-terminal
CC deletion NIK mutant proteins bind to IKK, thus inhibiting NIK/IKK
CC interaction
XX SQ Sequence 2844 BP; 665 A; 887 C; 809 G; 483 T; 0 U; 0 Other;

Query Match 37.1%; Score 977; DB 2; Length 2844;
Best Local Similarity 100.0%; Pred. No. 9.9e-262;
Matches 977; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCCTCTCAGCCAGGCGCATCCAGAGGGGCTGAGGAAGAGCCATCCAGCGCTGT 60
Db 1868 CCCCTCTCAGAGCCAGGCGCATCCAGAGGGGCTGAGGAAGAGCCATCCAGCGCTGT 1927
Qy 61 CTGACGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120
Db 1928 CTGACGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 1987
Qy 121 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCAATATACC 180
Db 1988 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCAATATACC 2047
Qy 181 ACCAGACCTCTCCATGCCAGCCGAGAGAGCTTTCGCCAAGGGGCCAGGGCCCCGGCCAG 240
Db 2048 ACCAGACCTCTCCATGCCAGCCGAGAGAGCTTTCGCCAAGGGGCCAGGGCCCCGGCCAG 2107

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:18:24 ; Search time 7480.59-Seconds
(without alignments)
17042.190 Million cell updates/sec

Title: US-09-155-676B-3
Perfect score: 2631
Sequence: 1 cccctctcacagccagcc.....acgtgacctcgaggggg 2631

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_on:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2553.8	97.1	2631	6	A66647
2	1851.2	70.4	4596	6	A66650 Sequence 3
3	1850.2	70.3	4596	6	AX429240 Sequence 6
4	1850.2	70.3	4596	6	AX774978 Sequence
5	1850.2	70.3	4596	9	HSNIK
6	1788.8	68.0	4452	9	BC035576
7	1781.8	67.7	4584	6	CQ723104 Sequence
8	1771.6	67.3	3152	6	CQ845924 Sequence
9	1771.6	67.3	3152	9	AK131438
10	1133.4	43.1	3156	6	AR062290 Sequence
11	1133.4	43.1	3156	6	AR063065 Sequence
12	1133.4	43.1	3156	6	AR068515 Sequence
13	1133.4	43.1	3156	6	BD062401 NIK prote
14	977	37.1	2844	6	AR429673 Sequence
15	973.4	37.0	48032	9	AC003963 Homo sapi
16	973.4	37.0	96875	9	AC008105 Homo sapi
17	973.4	37.0	100836	2	AC008105 Homo sapi
18	971	36.9	142326	2	AC024047 Homo sapi
19	672.8	25.6	2829	10	AF143094 Mus muscu

C 20	349.4	13.3	2475	6	AX834630
C 21	349.4	13.3	2475	9	AK097219 Homo sapi
C 22	283	10.8	337	11	G27543
C 23	275.8	10.5	162192	2	AC150796 Bos tauru
C 24	266.6	10.1	70748	2	AC137575 Homo sapi
C 25	238.4	9.1	698	11	BV166476
C 26	205.8	7.8	476	6	AX778201
C 27	188.8	7.2	213625	2	AC120950
C 28	188.8	7.2	223438	10	AL662804
C 29	188.8	7.2	239464	2	AC103079
C 30	138	5.2	70748	2	AC137575
C 31	100.6	3.8	101	11	BV184269
C 32	94.2	3.6	374	10	AY091549S1
C 33	91	3.5	907	9	BC031942
C 34	88.4	3.4	455	6	CQ490185
C 35	88.4	3.4	455	6	CQ496042
C 36	88.4	3.4	476	6	CQ480458
C 37	88.4	3.4	504	6	CQ501629
C 38	88.4	3.4	504	6	CQ510553
C 39	73.8	2.8	463	6	CQ471290
C 40	66	2.5	125020	9	AF429315
C 41	63.8	2.4	125020	9	AF429315 Homo sapi
C 42	60.4	2.3	111803	2	AC136172
C 43	58.2	2.2	2000	6	AX655393
C 44	53.2	2.0	1253	6	A66648
C 45	52.8	2.0	65	6	CQ556101

ALIGNMENTS

RESULT 1
LOCUS A66647 2631 bp DNA linear PAT 29-MAR-1999
DEFINITION Sequence 3 from Patent WO9737016.
ACCESSION A66647
VERSION A66647.1 GI:4538139
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2631)
Wallach,D., Malinin,N., Boldin,M., Kovalenko,A. and Mett,I.
MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR
PREPARATION AND USE
JOURNAL Patent: WO 9737016-A 3 09-OCT-1997;
YEDA RES & DEV (IL)
COMMENT Other publication AU 2175597 19971022.
FEATURES
Location/Qualifiers
source
1..2631
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 97.1%; Score 2553.8; DB 6; Length 2631;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCCCTCTCACAGCCCGAGGCGATCCAGAGGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
Db 1 CCCCTCTCACAGCCCGAGGCGATCCAGAGGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
Qy 61 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120
Db 61 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120
Qy 121 GCCCTTGGAGGGAGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCCAATTACC 180
Db 121 GCCCTTGGAGGGAGAGATATAAGAACCAAGACATCCACCGCCCAATTCAAGCCCAATTACC 180
Qy 181 ACCAGACCTTCATGCTCCAGCCGAGAGAGCTTTTCGCCAAGGGGCCCGAGGGCCCGGCCAG 240
Db 181 ACCAGACCTTCATGCTCCAGCCGAGAGAGCTTTTCGCCAAGGGGCCCGAGGGCCCGGCCAG 240

Db 181 ACCAGACCTCATGCCAGCCGAGAGAGCTTTGCCAGAGGGCCCGAGGGCCCGGCCAG 240
Qy 241 CTGAGAGACAACAGGCGAGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGAGCCCGAG 300
Db 241 CTGAGAGACAACAGGCGAGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGAGCCCGAG 300
Qy 301 AGCCAAAACAAGTCTCTCTCTTGAATTTGAGCAGAGGAGAGTCTGGGATGTGGAAACCT 360
Db 301 AGCCAAAACAAGTCTCTCTCTTGAATTTGAGCAGAGGAGAGTCTGGGATGTGGAAACCT 360
Qy 361 TACCTCTGTCCTCTGAGGCGAGCCCTTGCCAGAAACCCAGCTCCACAGAGCGGAAAG 420
Db 361 TACCTCTGTCCTCTGAGGCGAGCCCTTGCCAGAAACCCAGCTCCACAGAGCGGAAAG 420
Qy 421 CAACCGTCCCGAGCAGGAATTCGACGAGCTGGAATAAGATTAATTCCTCAACAGCCTGT 480
Db 421 CAACCGTCCCGAGCAGGAATTCGACGAGCTGGAATAAGATTAATTCCTCAACAGCCTGT 480
Qy 481 CCCAGCATTTTCTCTGGAGGAGGAGCAAAATTCCTCGTGCCTCAGCATCGACAGCC 540
Db 481 CCCAGCATTTTCTCTGGAGGAGGAGCAAAATTCCTCGTGCCTCAGCATCGACAGCC 540
Qy 541 TCTCCCTGTCGATGACAGTGAAGAAACCCATCAAGGCTCTCAAAAGCTCGCGGACA 600
Db 541 TCTCCCTGTCGATGACAGTGAAGAAACCCATCAAGGCTCTCAAAAGCTCGCGGACA 600
Qy 601 CCTGAGCTCAGCGTACACTCTGAGCAGCCAGCCGAGCTCGAAGCTCCAGCTGGA 660
Db 601 CCTGAGCTCAGCGTACACTCTGAGCAGCCAGCCGAGCTCGAAGCTCCAGCTGGA 660
Qy 661 ACATGTGTGCGCCGGGGGCGCCACCGACACCCCAAGCTATTCAATGGTGTGAAAG 720
Db 661 ACATGTGTGCGCCGGGGGCGCCACCGACACCCCAAGCTATTCAATGGTGTGAAAG 720
Qy 721 TCCAAATACAGTCTCTTAATGTGAAACACTGCAATCGGAGTTTCAACGGGTCAAAG 780
Db 721 TCCAAATACAGTCTCTTAATGTGAAACACTGCAATCGGAGTTTCAACGGGTCAAAG 780
Qy 781 TGGAGACATCGGCATCGGCATCAGCAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA 840
Db 781 TGGAGACATCGGCATCGGCATCAGCAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA 840
Qy 841 CCAAGACGGGCGAGCTGTCGTACGACATGAGGTGCCAGACTCGGCGATCGACCTGC 900
Db 841 CCAAGACGGGCGAGCTGTCGTACGACATGAGGTGCCAGACTCGGCGATCGACCTGC 900
Qy 901 AGTGCACTGGGCCCTGATGGCAGCTTGCCTTGGAGCTGGAGGTCAAGCATGGCCAGC 960
Db 901 AGTGCACTGGGCCCTGATGGCAGCTTGCCTTGGAGCTGGAGGTCAAGCATGGCCAGC 960
Qy 961 TGGAGACAGCCCTAACCTGCGCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT 1020
Db 961 TGGAGACAGCCCTAACCTGCGCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT 1020
Qy 1021 TCTGTCTGGTGCAAGTGTGCTTGAAGAAACAGGCTCAGCGTTCACAGGGGATYTG 1080
Db 1021 TCTGTCTGGTGCAAGTGTGCTTGAAGAAACAGGCTCAGCGTTCACAGGGGATYTG 1080
Qy 1081 NCCAGCCCCCGGCTCARGANTGGGAAACAGGGCTTCGNCAGNAGCNAAAGTNGGGG 1140
Db 1081 NCCAGCCCCCGGCTCARGANTGGGAAACAGGGCTTCGNCAGNAGCNAAAGTNGGGG 1140
Qy 1141 CAAGCAGAAATGCTCCAGGATTTACANCTGAGCCNCTGCCCCANCCCTGCTGAADA 1200
Db 1141 CAAGCAGAAATGCTCCAGGATTTACANCTGAGCCNCTGCCCCANCCCTGCTGAADA 1200
Qy 1201 AAAAAYTNCCGCCAGTGAAGAGACAGAGAGATGNCNAGGAGTTNNAACCTYGGGAA 1260
Db 1201 AAAAAYTNCCGCCAGTGAAGAGACAGAGAGATGNCNAGGAGTTNNAACCTYGGGAA 1260
Qy 1261 ACMAAACAGGATCTTTNTTCTGCCCTGCTCCAGTNCGAGTTGGCCTGNAACCCGCTTGG 1320
Db 1261 ACMAAACAGGATCTTTNTTCTGCCCTGCTCCAGTNCGAGTTGGCCTGNAACCCGCTTGG 1320

Qy 1321 ANTCACTGACCATTTGTTGGCAGANCAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGAAGG 1380
Db 1321 ANTCACTGACCATTTGTTGCGAGANCAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGAAGG 1380
Qy 1381 GTGGGCGAGCCCTTCGGCCCTCAACCTTNCAGAGGTCTGTGTCAGAGTCTCAAGTGTGTA 1440
Db 1381 GTGGGCGAGCCCTTCGGCCCTCAACCTTNCAGAGGTCTGTGTCAGAGTCTCAAGTGTGTA 1440
Qy 1441 AGGNCCTCAAACTCAGGNTTCAGTGCAGAACAGGTNCAGCAGGTATGCCCGCCGNTA 1500
Db 1441 AGGNCCTCAAACTCAGGNTTCAGTGCAGAACAGGTNCAGCAGGTATGCCCGCCGNTA 1500
Qy 1501 GGTAAANNNGGGGCCCTCTNAAACCCCTTGCCTTNGGCCCTNACCTTNGGCCAGCTCANCC 1560
Db 1501 GGTAAANNNGGGGCCCTCTNAAACCCCTTGCCTTNGGCCCTNACCTTNGGCCAGCTCANCC 1560
Qy 1561 CTTTGTGGGTAGGGGAAAAGATGCTGACCTCGGGAAGGCTTWCCTCGTAGAATAACAC 1620
Db 1561 CTTTGTGGGTAGGGGAAAAGATGCTGACCTCGGGAAGGCTTWCCTCGTAGAATAACAC 1620
Qy 1621 CACACTTTTCAGGTGTGTGCAACACAGGTCTCTGAGTTGACCTCTGCTGAGCCAGGACC 1680
Db 1621 CACACTTTTCAGGTGTGTGCAACACAGGTCTCTGAGTTGACCTCTGCTGAGCCAGGACC 1680
Qy 1681 AAAGAAGGTGTGAAGTGAAGTGTCTCAGTNCCTCAGTGCAGATGTGCCCCCTTTGCTGCTG 1740
Db 1681 AAAGAAGGTGTGAAGTGAAGTGTCTCAGTNCCTCAGTGCAGATGTGCCCCCTTTGCTGCTG 1740
Qy 1741 GCTACCACTCTTCCCAGAGCAGCCCGCCAGCCCTTTCAGGCCAGCAGCTGCCCCAG 1800
Db 1741 GCTACCACTCTTCCCAGAGCAGCCCGCCAGCCCTTTCAGGCCAGCAGCTGCCCCAG 1800
Qy 1801 ACTCGCTGGCACTCAGTTCCTCTCATCTGTAAGGTGAAGGTGATGAGGATATGCTGA 1860
Db 1801 ACTCGCTGGCACTCAGTTCCTCTCATCTGTAAGGTGAAGGTGATGAGGATATGCTGA 1860
Qy 1861 CAGAAACAGTCTGTGGATGGAACATGATGCTTNAAGNNAGCAGCAGAGAGAGCY 1920
Db 1861 CAGAAACAGTCTGTGGATGGAACATGATGCTTNAAGNNAGCAGCAGAGAGAGCY 1920
Qy 1921 TCCGGGCCCCAGNCCCACCTNATCAGTGTNCCAGCGTGTCTGCTGCTGCTGCTGCTG 1980
Db 1921 TCCGGGCCCCAGNCCCACCTNATCAGTGTNCCAGCGTGTCTGCTGCTGCTGCTGCTG 1980
Qy 1981 GCTNCAGNCACTANCACTGACACTNCACTTNGGCCCTGCCCCCANGAGGGTACTG 2040
Db 1981 GCTNCAGNCACTANCACTGACACTNCACTTNGGCCCTGCCCCCANGAGGGTACTG 2040
Qy 2041 CCGNACGGCACTTTGCACTCTGATGNACCTCAAGCACTTTTATGGCTNGCCCTCTNNG 2100
Db 2041 CCGNACGGCACTTTGCACTCTGATGNACCTCAAGCACTTTTATGGCTNGCCCTCTNNG 2100
Qy 2101 GCAGGNCAGGNCAGGNCAGTGAACCTGTAGNAGCATANGCAANGCAGGAGATGG 2160
Db 2101 GCAGGNCAGGNCAGGNCAGTGAACCTGTAGNAGCATANGCAANGCAGGAGATGG 2160
Qy 2161 GGTGNAAGGNCACAGTCTTGTGAGCTGTCAACATGCACTGTGACTNCTCAAACTCTTN 2220
Db 2161 GGTGNAAGGNCACAGTCTTGTGAGCTGTCAACATGCACTGTGACTNCTCAAACTCTTN 2220
Qy 2221 NCCAGNATTTCTTAAGNATAGNCCCTTNCCTTNGCCCTTCCAGCTTAGCCTCTTCT 2280
Db 2221 NCCAGNATTTCTTAAGNATAGNCCCTTNCCTTNGCCCTTCCAGCTTAGCCTCTTCT 2280
Qy 2281 CCCAGGGAGCTTANCTCAGGACTCACTGAGCATTAATCAGCTGTGNAATCGTCAGGGGG 2340
Db 2281 CCCAGGGAGCTTANCTCAGGACTCACTGAGCATTAATCAGCTGTGNAATCGTCAGGGGG 2340
Qy 2341 TGTCTGTAGCTCAACCTCTGGGGCAGGGGACCCCGAGACTCCGTGGGAGAGCTCAT 2400
Db 2341 TGTCTGTAGCTCAACCTCTGGGGCAGGGGACCCCGAGACTCCGTGGGAGAGCTCAT 2400

Qy 2401 TCCACATCTTGGCCAGACAGCCTTNGTCCAGCTGTCCACATTCAGTCACTGCTCCC 2460
Db 2401 TCCACATCTTGGCCAGACAGCCTTNGTCCAGCTGTCCACATTCAGTCACTGCTCCC 2460
Qy 2461 GGGAGAGAGCCCGGCCCCAGACACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTG 2520
Db 2461 GGGAGAGAGCCCGGCCCCAGACACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTG 2520
Qy 2521 GGTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGTGATGACAAATGTTAAAAAAG 2580
Db 2521 GGTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGTGATGACAAATGTTAAAAAAG 2580
Qy 2581 GCCTTCGTGGCCTCGAATCAAGCTTATCGATACCGTCGACCTCGAGGGGG 2631
Db 2581 GCCTTCGTGGCCTCGAATCAAGCTTATCGATACCGTCGACCTCGAGGGGG 2631

RESULT 2

LOCUS A66650 A66650 4596 bp DNA linear PAT 29-MAR-1999
DEFINITION Sequence 6 from Patent WO9737016.
ACCESSION A66650
VERSION A66650.1 GI:4538141

KEYWORDS
SOURCE unidentified
ORGANISM unidentified

REFERENCE 1 (bases 1 to 4596)
Wallach,D., Malinin,N., Boldin,M., Kovalenko,A. and Mett,I.
AUTHORS MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR
TITLE PREPARATION AND USE

JOURNAL Patent: WO 9737016-A 6 09-OCT-1997;

COMMENT YEDA RES & DEV (IL)

Other publication AU 2175597 19971022.

FEATURES

Location/Qualifiers
1..4596
source
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 70.4%; Score 1851.2; DB 6; Length 4596;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2335; Conservative 4; Mismatches 159; Indels 81; Gaps 34;
Qy 1 CCCCTCTCACAGCCAGCCATCCAGAGGGCTTGAGGAAGAGCCATCCACCGCGTGT 60
Db 2099 CCCCTCTCACAGCCAGCCATCCAGAGGGCTTGAGGAAGAGCCATCCACCGCGTGT 2158
Qy 61 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA 120
Db 2159 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA 2218
Qy 121 GCCCTTGGAGGGAGATATAAAGAACCAAGACATCCACCGCAAAATCAAGCCAAATACC 180
Db 2219 GCCCTTGGAGGGAGATATAAAGAACCAAGACATCCACCGCAAAATCAAGCCAAATACC 2278
Qy 181 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCGCCAAGGGCCCGAGGGCCCGGCCAG 240
Db 2279 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCGCCAAGGGCCCGAGGGCCCGGCCAG 2338
Qy 241 CTGAGGAGACAAGGAGAGCCCTTAAGTCTCCAGCCTCTCTCCACCAAGAGCCCGCCAG 300
Db 2339 CTGAGGAGACAAGGAGAGCCCTTAAGTCTCCAGCCTCTCTCCACCAAGAGCCCGCCAG 2398
Qy 301 AGCCAAACAAGTCTCTCCCTTGAATTTGAGCAAGAGGAGTCTGGGATGAGAAACCTT 360
Db 2399 AGCCAAACAAGTCTCTCCCTTGAATTTGAGCAAGAGGAGTCTGGGATGAGAAACCTT 2458
Qy 361 TACCTCTGTCTCTCCCTGGAGCCAGCCCTGCGCAAGAACCCAGCTCACAGAGCGGAAG 420
Db 2459 TACCTCTGTCTCTCCCTGGAGCCAGCCCTGCGCAAGAACCCAGCTCACAGAGCGGAAG 2518

Qy 421 CAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATTATTCTCAACAGCCTGT 480
Db 2519 CAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATTATTCTCAACAGCCTGT 2578
Qy 481 CCCAGCCATTTCTCTCGAGGAGCAGGAGAAATTTCTCTCGTGTCTCAGCATCGACAGCC 540
Db 2579 CCCAGCCATTTCTCTCGAGGAGCAGGAGGAAATTTCTCTCGTGTCTCAGCATCGACAGCC 2638
Qy 541 TCTCCCTGTCCGATGACAGTGAAGAAACCCATCAAAAGGCTCTCAAAAGCTTCGGGGACA 600
Db 2639 TCTCCCTGTCCGATGACAGTGAAGAAACCCATCAAAAGGCTCTCAAAAGCTTCGGGGACA 2698
Qy 601 CCCTGAGCTCAGGCGTACACTCTTGAGAGCAGCAGGCGGAGCTCGAAGCTCCAGCTGGA 660
Db 2699 CCCTGAGCTCAGGCGTACACTCTTGAGAGCAGCAGGCGGAGCTCGAAGCTCCAGCTGGA 2758
Qy 661 ACATGGTGTGGCCCGGGCGGCCACCGACACCCCAAGCTATTTCATGCTGTAAG 720
Db 2759 ACATGGTGTGGCCCGGGCGGCCACCGACACCCCAAGCTATTTCATGCTGTAAG 2818
Qy 721 TCCAAATACAGTCTCTTAATGGTGAAACACCTCGACATCCCGGGTCAAAG 780
Db 2819 TCCAAATACAGTCTCTTAATGGTGAAACACCTCGACATCCCGGGTCAAAG 2878
Qy 781 TGGGAGACATCGCCACTTGGCATCAGAGCCAGATCCAGCTGCGAGCTTCAAGCTTGTCA 840
Db 2879 TGGGAGACATCGCCACTTGGCATCAGAGCCAGATCCAGCTTCAAGCTTGTCA 2938
Qy 841 CCAAAGAGGGCAGCCCTTTCGCTACGACATGGAGGTGCCAGATTCGGGGCATCGACCTGC 900
Db 2939 CCAAAGAGGGCAGCCCTTTCGCTACGACATGGAGGTGCCAGATTCGGGGCATCGACCTGC 2998
Qy 901 AGTGACACTGGCCCTGTATGGCAGCTTCGCTTGGAGCTTGGAGGTCAAGCATGGCCAGC 960
Db 2999 AGTGACACTGGCCCTGTATGGCAGCTTCGCTTGGAGCTTGGAGGTCAAGCATGGCCAGC 3058
Qy 961 TGGAGAACAGCCCTAACCTCGCCCTCCACCGCGGTCCACATCGCCGGAAGAGAGCT 1020
Db 3059 TGGAGAACAGCCCTAACCTCGCCCTCCACCGCGGTCCACATCGCCGGAAGAGAGCT 3118
Qy 1021 TCCTGCTCGGTGACAGATGCTGCCCTGAAACACAGCTCAGCGCTTCCAGGGGATTTG 1080
Db 3119 TCCTGCTCGGTGACAGATGCTGCCCTGAAACACAGCTCAGCGCTTCCAGGGGATTTG 3176
Qy 1081 NCCAGCCCCCGGCTCARCAGNTGGGAACACAGGCGCTTCGNCAGCAGNAAGGTNGGGG 1140
Db 3177 GCCAGCCCCCGGCTCA--CAGTGGGAACACAGGCGCTCG-----CAGCAGCAAGGTGGG 3229
Qy 1141 CAAGCNAAGATGCTCCAGGATTTTCAANCTGAGCCCTGCCCCANCTGCTGTAADA 1200
Db 3230 GCAAGCAGAAATGCTCCAGGATTTTCACTGAGCCCTGCCCCAC-----CCTGCTGA 3283
Qy 1201 AAACATYNNCCGCCACGTGAAGAGACAGAGGATGNCAGGAGTTNNACCTTGGGGAA 1260
Db 3284 AAACATYNNCCGCCACGTGAAGAGACAGAGGATGNCAGGAGTTNNACCTTGGGGAA 3339
Qy 1261 ACAAACACAGGATCTTTNTTCTGCCCCCTGCTCCAGTTCGAGTTGGCTTGCCTGNAACCGCTGG 1320
Db 3340 ACAAACACAGGATCTTTT-TTCTGCCCCCTGCTCCAGT--CGAGTTGGCTGA--CCCGCTTG 3395
Qy 1321 ANTCAGTGACCAATTTGTTGGCAGNACAGGGAGAGAGCTTCCAGCTTGGGTCAAGAGGG 1380
Db 3396 GATCAGTGAACCAATTTGTTGGCAGA-CAGGGAGAGAGAGCTTCCAGCTTGGGTCAAGAGGG 3454
Qy 1381 GTGGGCGAGCCCTTCGGCCCTTCACTTNCAGGCTCTGTGNAAGAGTGTCAAGTGTGA 1440
Db 3455 GTGGGCGAGCCCTTCGGCCCTTCACTTNCAGGCTCTGTGNAAGAGTGTCAAGTGTGA 3512
Qy 1441 AGGNCACCAANCTCAGGNTTCAGTGCAGAACACAGGTNCAGCAGGTATGCCCGCCCGNTA 1500
Db 3513 AGGNCACCAANCTC---AGGTTTCAGTGCAGAACACAGGT--CAGCAGGTATGCCCGCCCGTAG 3568
Qy 1501 GGTTAANNNGGGGCCCTCTNAAACCCCTTGGCTTNGGCTTNCACCTNGGCCAGCTCANCC 1560

Db 3569 GTTAA----GGGGCCCTCTAAACCCCTTGCTGGCCTCA---CCTGCCAGTCTCA-CCC 3620
Qy 1561 CTTTGGGTGTAGGGGAAAGAAATGCTGACCCCTGGGAAGGCTWCCTGCTAGATAACAC 1620
Db 3621 CTTTGGGTGTAGGGGAAAGAAATGCTGACCCCTGGGAAGGCT-CCCTGGTAGAATAACAC 3679
Qy 1621 CACACTTTTCAGTGTGTTGCAACACAGGTCTGTGAGTTGACCTCTGTGTTTTCAGCCAAAGACC 1680
Db 3680 CACACTTTTCAGTGTGTTGCAACACAGGTCTGTGAGTTGACCTCTGTGTTTTCAGCCAAAGACC 3739
Qy 1681 AAAGAAGGTGTAAAGTGAAGTGGTTCAGTGNCCCAAGACATGTGCCCTTTGCTGCTG 1740
Db 3740 AAAGAAGGTGTAAAGTGAAGTGGTTCAGT-CCCCAGACATGTGCCCTTTGCTGCTG 3798
Qy 1741 GCTACCACTTTTCCCGAGGACAGGCCCCGAGGCCCTTCAGGCCCAAGCACTGCCCCAG 1800
Db 3799 GCTACCACTTTTCCCGAGGACAGGCCCCGAGGCCCTTCAGGCCCAAGCACTGCCCCAG 3858
Qy 1801 ACTCGCTGGCACTCAGTTCCTCATCTGTAAAGTGAAGGTTGATGCAAGTATGCTGTA 1860
Db 3859 ACTCGCTGGCACTCAGTTCCTCATCTGTAAAGTGAAGGTTGATGCAAGTATGCTGTA 3918
Qy 1861 CAGGAACAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920
Db 3919 CAGGAACAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3975
Qy 1921 TCCGGCGCCCCAGNCCCACTNATCAGTGTNCCAGCGTGTCTGTTNCCCAAGNAGACA 1980
Db 3976 TCCGGCGCCCCAGNCCCACTNATCAGTGT---CCAGCGTGTGTTTCCCGAGGACACAG- 4031
Qy 1981 GCTNCAAGNCACTNCACTGACATNCACTGACATNCACTGACATNCACTGACATNCACTG 2040
Db 4032 -----TCAGCATCACATGACATCACATGACATCACATGACATCACATGACATCACAT 4082
Qy 2041 CCGNACGGCACTTTCAGTCTGATGACATGACATGACATGACATGACATGACATGACATGAC 2100
Db 4083 CCG-ACGGCACTTTCAGTCTGATGACATGACATGACATGACATGACATGACATGACATGAC 4139
Qy 2101 GCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGG 2160
Db 4140 GGGCAGGCGAGGCGAGTGACACT-----GTAGGACATAGCAAGCCAGGAGATGG 4189
Qy 2161 GGTGNAAGGNAACAGTCTTGAGTGTCCAGCATGATGATGATGATGATGATGATGATGATGAT 2220
Db 4190 GGTGNA--GGGACACAGTCTTGAGCTGTCCA-CATGATGTGACT-CTTCAAACCTCTT- 4244
Qy 2221 NCCAGNATTTCTTAAGAATAGCAGNCCCTTNCCTTNCCTTNCCTTNCCTTNCCTTNCCTT 2280
Db 4245 --CCAGATTTCTTAAGATAGCAGNCCCTTNCCTTNCCTTNCCTTNCCTTNCCTTNCCTT 4300
Qy 2281 CCCAGGGAGTANCTCAGGACTCACTGATGATGATGATGATGATGATGATGATGATGATGATG 2340
Db 4301 CCCAGGGAGGCTA-CTCAGGACTCAGTATGATGATGATGATGATGATGATGATGATGATGATG 4358
Qy 2341 TGTCTGTAGCTCAACTCTCTGGGCGAGGGGAGCGCGAGACTCCGTGGGAGAGTCTCAT 2400
Db 4359 TGTCTGTAGCTCAACTCTCTGGGCGAGGGGAGCGCGAGACTCCGTGGGAGAGTCTCAT 4418
Qy 2401 TCCCACTTTTCCCAAGACAGCTTTTNGTCCAGTGTCCCATGATGATGATGATGATGATGATG 2460
Db 4419 TCCCACTTTTCCCAAGACAGCTTTT-GTCCAGTGTCCCATGATGATGATGATGATGATGATG 4477
Qy 2461 GGGGAGAGAGCCCGGCCCCCAGACATAGAGACTCAGCTTGGTACTGACAGTCTG 2520
Db 4478 GGGGAGAGAGCCCGGCCCCCAGACATAGAGACTCAGCTTGGTACTGACAGTCTG 4537
Qy 2521 GGTGTGTAGAGACTCTTTGTAAAGCAATAAAGTTTGGGTGATGACAAATGTTAAAAAAA 2579
Db 4538 GGTGTGTAGAGACTCTTTGTAAAGCAATAAAGTTTGGGTGATGACAAATGTTAAAAAAA 4596

AX429240 LOCUS AX429240 4596 bp DNA linear PAT 21-JUN-2002
DEFINITION Sequence 17 from Patent EP1201765.
ACCESSION AX429240
VERSION AX429240.1 GI:21540550
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Schubart,D., Habenberger,P., Stein-Gerlach,M. and Bevec,D.
TITLE Cellular kinases involved in cytomegalovirus infection and their inhibition
JOURNAL Patent: EP 1201765-A 17 02-MAY-2002;
Axxima Pharmaceuticals Aktiengesellschaft (DE)
FEATURES
source location/Qualifiers
1. .4596
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="N/A"

ORIGIN

Query Match 70.3%; Score 1850.2; DB 6; Length 4596;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

Qy 1 CCCTCTCACAGCCAGGCCATCCAGAGGGGTGAGGAAGAGCCATCCACCGGTGT 60
Db 2100 CCCTCTCACAGCCAGGCCATCCAGAGGGGTGAGGAAGAGCCATCCACCGGTGT 2159
Qy 61 CTGAGCGGAGCTGGGAGGAAGGTGAACCGGGGCACTACAGCAAGTGGGAGTCTGAAGA 120
Db 2160 CTGAGCGGAGCTGGGAGGAAGGTGAACCGGGGCACTACAGCAAGTGGGAGTCTGAAGA 2219
Qy 121 GCCCTTGGAGGGGAGATATAAGAACCAAGACATCCACCGCAAAATCAAGCCAAATTACC 180
Db 2220 GCCCTTGGAGGGGAGATATAAGAACCAAGACATCCACCGCAAAATCAAGCCAAATTACC 2279
Qy 181 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTGCGCAAGGGGCCCHAGGGCCCGGCGCAG 240
Db 2280 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTGCGCAAGGGGCCCHAGGGCCCGGCGCAG 2339
Qy 241 CTGAGGAGACAAAGGAGAGCCCTTAAGTCCAGGCTCTCTCCACACAGAGCCCGCAG 300
Db 2340 CTGAGGAGACAAAGGAGAGCCCTTAAGTCCAGGCTCTCTCTCCACACAGAGCCCGCAG 2399
Qy 301 AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGAAACCT 360
Db 2400 AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGAAACCT 2459
Qy 361 TACCT 420
Db 2460 TACCT 2519
Qy 421 CAAACGTCCTGGAGCAGAACTGCGAGAGCTGGAATAGAATTAATTCCTCAACAGCCTGT 480
Db 2520 CAAACGTCCTGGAGCAGAACTGCGAGAGCTGGAATAGAATTAATTCCTCAACAGCCTGT 2579
Qy 481 CCAGCCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTGCTCTAGCATCGACGCC 540
Db 2580 CCAGCCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTGCTCTAGCATCGACGCC 2639
Qy 541 TCTCCCTGTCTGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCGCGGACA 600
Db 2640 TCTCCCTGTCTGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCGCGGACA 2699
Qy 601 CCCTGAGCTCAGCGTACACTCTCTGGAGCAGCAGCGCGAGGCTCGAAGCTCCAGTGA 660
Db 2700 CCCTGAGCTCAGCGTACACTCTCTGGAGCAGCAGCGCGAGGCTCGAAGCTCCAGTGA 2759
Qy 661 ACATGTGTGCGCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTGTGAAG 720

2760	Db	ACATGTGTCTCGCGCGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG	2819
721	Qy	TCCAAATACAGTCTCTTAATGGTGAACAACCTGCACATCCGGAGCTTCCACCGGTCAAAG	780
2820	Db	TCCAAATACAGTCTCTTAATGGTGAACAACCTGCACATCCGGAGCTTCCACCGGTCAAAG	2879
781	Qy	TGGGAGACATCGCCACTGTGGCATACGACAGCCAGATCCAGCTGCAGCCTTTCAGCTTGTGTC	840
2880	Db	TGGGAGACATCGCCACTGTGGCATACGACAGCCAGATCCAGCTGCAGCCTTTCAGCTTGTGTC	2939
841	Qy	CCAAAGACGGGACGCTGTTCGCTACGACATGGAGGTGCCAGACTCGGGGATCGACCTGC	900
2940	Db	CCAAAGACGGGACGCTGTTCGCTACGACATGGAGGTGCCAGACTCGGGGATCGACCTGC	2999
901	Qy	AGTGCACATCGGCCCTCGATGGCAGCTTTCGCTCGAGCTGGAGGTCAAGCATGCGCCAGC	960
3000	Db	AGTGCACATCGGCCCTCGATGGCAGCTTTCGCTCGAGCTGGAGGTCAAGCATGCGCCAGC	3059
961	Qy	TGGAGAACAGGCGCTTAACCTTCGCCCTCCACCGCGGCTCCACATGCGGAAAGACGCT	1020
3060	Db	TGGAGAACAGGCGCTTAACCTTCGCCCTCCACCGCGGCTCCACATGCGGAAAGACGCT	3119
1021	Qy	TCCCTGCTCGGTGCAGATGCTCCCTGAAACACAGGCTCAGCGTTCCACGGGATTTG	1080
3120	Db	TCCCTGCTCGGTGCAGATGCTCCCTGAAACACAGGCTCAGCGTTCCACGGGATTTG	3177
1081	Qy	NCCAGCCCCCGGCTCARGAGTGGGAAACAGGGCTTCGNCAGCNAGNAGGTTGGGG	1140
3178	Db	GCCAGCCCCCGGCTCA--CNGTGGGAACAGGGCTTCG-----CAGCAGCAAGTGGG	3230
1141	Qy	CAAGCNAGATGCTCCCGAGATTTCAACCTGAGCCCTGAGCCCTGCCCCANCCCTGCTGA	1200
3231	Db	GCAAGCAGATGCTCCCGAGATTTCAACCTGAGCCCTGCCCCAC-----CCTGCTGA	3284
1201	Qy	AACAYTNCGCCACGTTGAAGACAGAGAGAGAGATGNCAGGATGNNACCTTGGGAA	1260
3285	Db	AAAACATCGCCACGTTGAAGACAGAGAGAGAGATGNCAGGAGGATGNC-----AGGAGTTACCTGGGAA	3340
1261	Qy	ACAAAACAGGATCTTTNTTCTGCCCTGCTCCAGTNCAGTGTGGCTGCTGACCCGCTTG	1320
3341	Db	ACAAAACAGGATCTTTTCTGCCCTGCTCCAGT--CGAGTTGGCCCTGA--CCCGCTTG	3396
1321	Qy	ANTCAGTGACCATTTGTTGGCAGANCAGGGAGAGCAGCTTCCAGCCTGGGTGAGAGGG	1380
3397	Db	GATCAGTGACCATTTGTTGGCAGA-CAGGGAGAGCAGCTTCCAGCCTGGGTGAGAGGG	3455
1381	Qy	GTGGGCAGCCCTTCGGCCCTCACCTTNCAGGCTGCTGTGAGAGTGTCAAGTGTGTA	1440
3456	Db	GTGGGCAGCCCTTCGGCCCTCACCTTNCAGGCTGCTGTGAGAGTGTCAAGTGTGTA	3513
1441	Qy	AGGNCCTCAANCTCAGNTTTCAGTGAGAACACAGTNCAGCAGTATGCCGCCGNTA	1500
3514	Db	AGGCCCTCAATC---AGGTTTCAGTGAGAACACAGT--CAGCAGGTATGCCGCCGNTAG	3569
1501	Qy	GGTTAANNNGGGCCCTCTNAACCCCTTGCTTNGGCTTNCACCTTNGGCCAGCTCANCCC	1560
3570	Db	GTTAA----GGGGCCCTCTAAACCCCTTGCTTGGCTCA--CCTGSCAGCTCA--CCC	3621
1561	Qy	CTTTTGGGTGAGGGGAAAAGATGCTCAGCCCTGGGAAGGCTTCCCTGTGATAGATAC	1620
3622	Db	CTTTTGGGTGAGGGGAAAAGATGCTCAGCCCTGGGAAGGCT--CCCTGTGATAGATAC	3680
1621	Qy	CACACTTTTCAGGTTGTTGAAACACAGTCTCAGTTGACTCTGTTTCAGCAAGGACC	1680
3681	Db	CACACTTTTCAGGTTGTTGAAACACAGTCTCAGTTGACTCTGTTTCAGCAAGGACC	3740
1681	Qy	AAAGAAGTGTGAAGTGAAGTGTCTCAGTNCCTCAGCATGTGCCCTTTGCTGCTG	1740
3741	Db	AAAGAAGTGTGAAGTGAAGTGTCTCAGT--CCCAGACATGTGCCCTTTGCTGCTG	3799
1741	Qy	GCTTACCACTCTTCCCCAGAGCAGCAGCCCGAGCCCTTTCAGGCTCAGCATGCTGCCAG	1800
3800	Db	GCTTACCACTCTTCCCCAGAGCAGCAGCCCGAGCCCTTTCAGGCTCAGCATGCTGCCAG	3859

Qy	1801	ACTCGTGGCACTCAGTTTCCCTCATCTGTTAAAGGTGAGGGTGATGCAGGATATGCCTGA	1860
Db	3860	ACTCGTGGCACTCAGTTTCCCTCATCTGTTAAAGGTGAGGGTGATGCAGGATATGCCTGA	3919
Qy	1861	CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNNAAGGNAAGCAGCAGAGAGACGY	1920
Db	3920	CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAA--GGAAAGCAGCAGAGAGACG-	3976
Qy	1921	TCGGCGCCCCCAGNCCCCCACTNATCAGTGTNCCAGCGTGTCTNGTGTTNCCCAGNAGCACA	1980
Db	3977	TCGGCGCCCCCAGCCCCCACTATCAGTGT---CCAGCGTGTGTTTCCCCAGAGCACAGC-	4032
Qy	1981	GCTNCAAGNATCANCACTGACACTNCAACCTNCGCTGCCCTNCGGCANGAGAGGCTACTG	2040
Db	4033	-----TCAGCATCACACTGACACTCACCTGCCCTGCCCTCGG--CCAGAGGGGTACTG	4083
Qy	2041	CCGNA CGGCACCTTTGCACNCTCTGATGNACCTCAAGACACTTTCATGGCTNGCCCTCTNNG	2100
Db	4084	CCG-ACGGCACTTTGCACT--CTGATGACCTCAAGACACTTTTCATGGCTGCCCTCTGGCA	4140
Qy	2101	GCAGGNCAGGNCAGGNCAGGNCAGTGCACNCTGTAGGNAGCATANGCAANGCCAGGAGATGG	2160
Db	4141	GGGCAGGCGAGGCGAGTGACACT-----GTAGGAGCATAGCAAGCCAGGAGATGG	4190
Qy	2161	GGTGNAAGGAGNACACAGTCTTTGAGCTGTGCANCATGATGTGATCTNCCCTCAAACTCTTN	2220
Db	4191	GGTGAA--GGGACACAGTCTTTGAGCTGTCCA-CATGCACTGTGACT--CCTCAAACTCTT-	4245
Qy	2221	NCCAGNATTTTCTTAAGATAGCANCCCCCTTNCCTTNGCCCATTTGCCCACTTAGCCTCTTCT	2280
Db	4246	--CCAGATTTCTCTAAGAATAGCACCCCC--TTCCCCCATTTGCCCACTTAGCCTCTTCT	4301
Qy	2281	CCCAGGGGAGCTANCTCAGGACCTCAGGTAGCATTAATCAGCTGTGTNAAATCGTCAGGGG	2340
Db	4302	CCCAGGGGAGCTA-CTCAGGACTCAGTAGCATTAATCAGCTGTG-AATCGTCAGGGGG	4359
Qy	2341	TGTTCTGTAGCTCAACCTCTCGGGCAGGGGAGCGCAGACTCCGTGGGAGAAGCTCAT	2400
Db	4360	TGTTCTGTAGCTCAACCTCTCGGGCAGGGGAGCGCAGACTCCGTGGGAGAAGCTCAT	4419
Qy	2401	TCCACATCTTGGCAAGACAGCCTTTNGTCCAGCTGTCCACATTTGAGTCAGACTGCTCCC	2460
Db	4420	TCCACATCTTGGCAAGACAGCCTTT-GTCCAGCTGTCCACATTTGAGTCAGACTGCTCCC	4478
Qy	2461	GGGGAGAGACCCCGGCCCCCAGCACATAAAGAACTGCAGCCTTGTCTATGCAGAGTCTG	2520
Db	4479	GGGGAGAGACCCCGGCCCCCAGCACATAAAGAACTGCAGCCTTGTCTATGCAGAGTCTG	4538
Qy	2521	GGTTGTAGAACTCTTTGTAAAGATAAAGTTTGGGGTGATGACAAATGTTAAAAA	2578
Db	4539	GGTTGTAGAACTCTTTGTAAAGATAAAGTTTGGGGTGATGACAAATGTTAAAAA	4596

RESULT 4

AX774978

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS
SOURCE

SOURCE ORGANISMS

STATUS

REFERENCE

AUTHORS

TITLE

JOURNAL

222

FEATURES

1000

/mol_type="unassigned DNA" /db_xref="taxon:9606"		Db		3060		TGAGAAACAGGCGCTAACCTGCGCTCCACCGCGCTCCACACTGCGCGAAAGCAGCCT		3119	
Query Match		Qy		1021		TCCTGCTGGGTGACGATGCTGCTGAAAAACACAGGCTCAGCGCTTCCAGGGGATGTG		1080	
Best Local Similarity		Db		3120		TCCTGCTGGGTGACGATGCTGCTGAAAAACACAGGCTCAGCGCTTCCAGGGGATGTG		3177	
Matches 2334; Conservative		Qy		1081		NCCAGCCCCCGCTCARGNATGGGAACACAGGCGCTCGNCAGCNAGNNAAGTNGGGG		1140	
		Db		3178		GCAGAGCCCCCGCTCA--CAGTGGGAACAGGGGCTCG-----CAGCAGAGGTGGG		3230	
		Qy		1141		CAAGCNAGATGCTCCAGGATTTTCAACNCTGAGCCNCTGCCCCCANCCCTGCTCAADA		1200	
		Db		3231		GCAAGCAGATGCTCCAGGATTTTCAACCTGAGCCCTGCCAC-----CCTGCTGA		3284	
		Qy		1201		AAACATYTCGCGCACGTGAAGACAGAGGAGGATGNCAGGATTNNAACCTTGGGGAA		1260	
		Db		3285		AAAAACATCGCCACGTGAAGACAGAGGAGGATGGC-----AGGATTTACCTGGGGAA		3340	
		Qy		1261		ACAAACAGGAGTCTTTTCTGCTCCCTGCTCCAGTNCAGATTTGGCTGNAACCGTTGG		1320	
		Db		3341		ACAAAACAGGATCTTTTCTGCTCCCTGCTCCAGT--CGAGTTGGCTGA--CCCGCTTG		3396	
		Qy		1321		ANTCAGTGACATTTGTTGGCAGANCAAGGGGAGAGAGCTTCCAGCTTGGGTGAGAGGG		1380	
		Db		3397		GATCAGTGACATTTGTTGGCAGA--CAGGGGAGAGCAGCTTCCAGCTTGGGTGAGAGGG		3455	
		Qy		1381		GTGGGCGAGCCCTCGGCGCTTCAACCTTCCAGCTGCTGTCGTCGTCGTCGTCGTCGTCG		1440	
		Db		3456		GTGGGCGAGCCCTTGGCGCTTCAACCT--CAGGCTGCTG--AGAGTGTCAAGTGTGTA		3513	
		Qy		1441		AGGNCACAAANTCAGGNTTCAGTGCAAAACAGGTCNAGCAGAGGTATGCCCGCCGNTA		1500	
		Db		3514		AGGCCCCAAATC---AGGTTCAAGTGCAAAACAGGT--CAGCAGGTATGCCCGCCGTAG		3569	
		Qy		1501		GTTTANNNGGGGCGCTCTTNAACCCCTTGCCTNGGCTTNCACCTNGGCGAGCTCANCCT		1560	
		Db		3570		GTTAA---GGGGCGCTCTTAAACCCCTTGCCTGCGCTCA---CCTGGCCAGCTCA---CCC		3621	
		Qy		1561		CTTTTGGGTGAGGGGAAAAGATGCTGACCTTGGGAAGGCTTCCCTTGGTGAATACAC		1620	
		Db		3622		CTTTTGGGTGAGGGGAAAAGATGCTGACCTTGGGAAGGCT--CCCTGTGTAATACAC		3680	
		Qy		1621		CACATTTTCAGGTTGTTGCAACAGAGTCTGAGTTGACCTCTGTTTCCAGCCAGGACC		1680	
		Db		3681		CACATTTTCAGGTTGTTGCAACAGAGTCTGAGTTGACCTCTGTTTCCAGCCAGGACC		3740	
		Qy		1681		AAAGAGGTGTTAAGTCAAGTGTCTCAGTNCACAGACATGTGCCCTTTTGTGCTG		1740	
		Db		3741		AAAGAGGTGTTAAGTCAAGTGTCTCAGT--CCCCAGACATGTGCCCTTTGCTGCTG		3799	
		Qy		1741		GCTACACTCTTTCCCGAGCAGCAGGCCCCGAGCCCTTCAAGGCCAGCATGTGCCCCAG		1800	
		Db		3800		GCTACACTCTTTCCCGAGCAGCAGGCCCCGAGCCCTTCAAGGCCAGCATGTGCCCCAG		3859	
		Qy		1801		ACTGCTGCACTCAATGCTTCCCTCATCTGTAAGGTGAAGGTGATGCAAGGATATGCTGA		1860	
		Db		3860		ACTGCTGCACTCAATGCTTCCCTCATCTGTAAGGTGAAGGTGATGCAAGGATATGCTGA		3919	
		Qy		1861		CAGGAACGTCTGTGATGATGATCATCATGTCCTTAAAGNAAAGCAGCAGAGAGAGAG		1920	
		Db		3920		CAGGAACGTCTGTGATGATGATCATCATGTCCTTAA--GGAAAGCAGCAGAGAGAGAG		3976	
		Qy		1921		TCGGGCGCCCCAGNCCCCTTATCAGTGTNCCAGGCTGCTGCTTCCCGAGAGCAGC-		4032	
		Db		3977		TCGGGCGCCCCAGNCCCCTTATCAGTGT---CCAGGCTGCTGTTTCCCGAGAGCAGC-		4083	
		Qy		1981		GCTNAGNATCANCACTGACACTNCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG		2040	
		Db		4033		-----TCAGCATCACACTGACACTCACCTTCCCTGCTGCTGCTGCTGCTGCTGCTG		4083	
		Qy		2041		CCGAGCGCACTTTCAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG		2100	

```
Db 4084 CCG-ACGGCACTTTGCACT--CTGATGACCTCAAGACACTTTTCATGGCTGCCCTCTTGCA 4140
Qy 2101 GCAGGNCAGGNCAGGNCAGTGCACACTGTAGNAGCATANGCAANGCCAGGAGATGG 2160
Db 4141 GGGCAGGCGAGGCGAGTGCACACT-----GTAGGAGCATAGCAAGCCAGGAGATGG 4190
Qy 2161 GGTGNAAGGNCAGACTCTTGAGCTGTCCANCATGTCATGTGACTNCTCAAACTCTTN 2220
Db 4191 GGTGAA--GGGACACAGCTTTGAGCTGTCCA-CATGATGTGACT-CCTCAAACTCTT- 4245
Qy 2221 NCCAGNATTTCTTAAGAAATAGCACCCTTTCCTCCCATTCCTCCAGCTTAGCTCTTCT 2280
Db 4246 --CCAGATTTCTTAAGAAATAGCACCTCC--TTCCCATTCCTCCAGCTTAGCTCTTCT 4301
Qy 2281 CCCAGGGAGTANTCTCAGSACTCAGCTAGCATTAATCAGCTGTGNAATCGTCAGGGGG 2340
Db 4302 CCCAGGGAGCTA-CTCAGGACTCAGCTAGCATTAATCAGCTGTG-AATGTCAGGGGG 4359
Qy 2341 TGTCTGTAGCTCAACTCTCTGGGCGAGGGAGCCGAGACTCTCGTGGGAGAGCTCAT 2400
Db 4360 TGTCTGTAGCTCAAACTCTCTGGGCGAGGGAGCCGAGACTCTCGTGGGAGAGCTCAT 4419
Qy 2401 TCCACACTCTTGGCAAGACAGCTTTNGTCCAGCTGTCCACATTTGAGTCAGACTGCTCCC 2460
Db 4420 TCCACACTCTTGGCAAGACAGCTTT-TGCCAGCTGTCCACATTTGAGTCAGACTGCTCCC 4478
Qy 2461 GGGGAGAGAGCCCGGCGCCAGCACATAAGAACTGCAGCTTTGGTACTGCAGACTCTG 2520
Db 4479 GGGGAGAGAGCCCGGCGCCAGCACATAAGAACTGCAGCTTTGGTACTGCAGACTCTG 4538
Qy 2521 GGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTGGGTGATGACAAATGTTAAAAA 2578
Db 4539 GGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTGGGTGATGACAAATGTTAAAAA 4596

RESULT 5
HSNIK
LOCUS HSNIK 4596 bp mRNA linear PRI 13-JAN-1998
DEFINITION H.sapiens mRNA for serine/threonine protein kinase, NIK.
ACCESSION Y10256
VERSION Y10256.1 GI:1841433
KEYWORDS MAP kinase; NIK protein; serine/threonine protein kinase.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Malinin,N.L., Boldin,M.P., Kovalenko,A.V. and Wallach,D.
MAP3K-related kinase involved in NF-kappaB induction by TNF, CD95
and IL-1
JOURNAL Nature 385 (6616), 540-544 (1997)
MEDLINE 97172277
PUBMED 9020361
2 (bases 1 to 4596)
Wallach,D.
Direct Submission
Submitted (23-DEC-1996) D. Wallach, The Weizmann Institute, Dept of
Membrane Research & Biophysics, Rehovot 76100, ISRAEL
NIK is a serine/threonine protein-kinase, resembling several MAP
kinase kinase kinases (MAP3K), that binds specifically to TRAF2, an
adaptor proteins associated, either directly or through interaction
with other adaptor proteins, with several receptors of the TNF/NGF
family. NIK overexpression in cells activates the transcription
factor NF kappa B. Cellular expression of kinase-deficient
NIK-mutants blocks NF kappa B induction by TNF, by either of the
two TNF receptors, by CD95 (Fas/Apo-1) and by TRADD, RIP and
MORT1/FADD, adaptor proteins that bind to these receptors. It also
blocks NF kappa B induction by IL-1.
FEATURES
Location/Qualifiers
1..4596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
```

```
misc_feature
137..149
/notes="lysine-rich region"
CDS
233..3076
/codon_start=1
/product="NIK, serine/threonine protein-kinase"
/protein_id="CAA71306.1"
/db_xref="GI:1841434"
/db_xref="GOA:Q99558"
/t_xref="UniProt/Swiss-Prot:Q99558"
```

```

/translation="MAVEMACPCAGSVAQOQKELPKPEKTPPLGKKQSSVVKLEA
VERSPVFCGWELINDVITKTAGESEAGPAALISIIQAECENSQSPSPFSPSRIFI
AGSKYQSQSSLSDOIENNVAHTEGWARVCMKGRKSKARKKKKKSKSLAHAGVA
LAKPLRTPQESCTIPQVEDSPGLGAPVYRNTQFTKPLKEPLGLQCFKQEGEGLR
LAPRSLHLKLI19LQCLNHVWLKHPQDGGPLPLPHTPPLPFPFPHLPQPKW
PHYLESPLKGLACVDSPLKSLACVDSPLKPLGPHLPSCLSRGAHEKFSVE
EHLVHALQGSVSSQASHLSLAKTWAAGRSRSPKPTEDNEGVLITKLPVDVE
YREVWATHQLRLGRGSFGEVHRMEDKQTGFQCAVKVRLVLEFRAELMACAGLTSP
RI19VLYGAVREGPWNI19FMBELGGSLQVLKEQCLPREDALYYLGOALEGLYLHS
RRI19HGVDKADNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDI19PTETMAPE
VILGRSDAKVDVMSCCMMLHMLNGCPHWTFQFRGPLCLKIASVEPPVRI19PSCAP
LTAQAIQELGRKEPIHRSVAELGGKVNRA19QVGLKSPWRGVEKPRHPPPNOAY
HQT19HAQPRELS19RAPRPAEETTGRAKLOPLPPEPPENKSP19LTLKESGMW
EPL19SLEPAPARNPSPSKATVPQELQQLEIEFLNSLSQPFSLEROBQILSCL
SIDLSLSDSEKNPSKASQSRD19LSSVHSSWQAERSSWMVLARGRPTDPS
YFNGVKVQIQLNGEHLHIREFHRVKVGD19ATG19SSQIPAAAFSLVTKDQGPVRYDME
VPDSDGLQCTLAPDGSFANRVRVHQLENRP"
```

```
misc_feature
401..653
/notes="kinase domain"
```

```
misc_feature
711..729
/notes="proline-rich region"
```

ORIGIN

```
Query Match 70.3%; Score 1850.2; DB 9; Length 4596;
Best Local Similarity 90.5%; Pred No. 0;
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;
Qy 1 CCCCTCTCACAGCCCGAGCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
Db 2100 CCCCTCTCACAGCCCGAGCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 2159
Qy 61 CTGAGCGGAGCTGGGAGGAAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120
Db 2160 CTGAGCGGAGCTGGGAGGAAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 2219
Qy 121 GCCCTTGGAGGGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCCAATACC 180
Db 2220 GCCCTTGGAGGGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCCAATACC 2279
Qy 181 ACCAGACCTTCCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCCCGGGCCAG 240
Db 2280 ACCAGACCTTCCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCCCGGGCCAG 2339
Qy 241 CTGAGGAGACAAGCGAGAGCCCTTAAGTCTCAGCTCTCTTCCACAGAGCCCCCAG 300
Db 2340 CTGAGGAGACAAGCGAGAGCCCTTAAGTCTCAGCTCTCTTCCACAGAGCCCCCAG 2399
Qy 301 AGCCAAACAGTCTCTCCCTTTCATTTTGACGAGGAGGAGTCTGGGATGTGGAAACCT 360
Db 2400 AGCCAAACAGTCTCTCTCCCTTTCATTTTGACGAGGAGGAGTCTGGGATGTGGAAACCT 2459
Qy 361 TACCTCTGTCTCTCTCGGAGCCAGCCCTCTCCAGAAACCCAGCTCACAGAGCGGAAAG 420
Db 2460 TACCTCTGTCTCTCTCGGAGCCAGCCCTCTCCAGAAACCCAGCTCACAGAGCGGAAAG 2519
Qy 421 CAACCGTCCCGAGCAGGAACTGCAGAGCTGGAAATAGAAATATTCTCTCAACAGCCTGT 480
Db 2520 CAACCGTCCCGAGCAGGAACTGCAGAGCTGGAAATAGAAATATTCTCTCAACAGCCTGT 2579
Qy 481 CCCAGCCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTCCTCAGCATCCAGACCC 540
Db 2580 CCCAGCCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTCCTCAGCATCCAGACCC 2639
```

Qy	541	TCTCCCTGTGCGATGACAGTGGAGAACAACCATTCAAAGCCTCTCAAAAGCTCGCGGACA	600
Db	2640	TCTCCCTGTGCGATGACAGTGGAGAACAACCATTCAAAGCCTCTCAAAAGCTCGCGGACA	2699
Qy	601	CCTTGAGCTCAGCGGTACACTCTCTGAGCAGCCAGGCCTCGAAGCTCCAGCTCGA	660
Db	2700	CCTTGAGCTCAGCGGTACACTCTCTGAGCAGCCAGGCCTCGAAGCTCCAGCTCGA	2759
Qy	661	ACATGTGTGCGCCCGGGGGCGGCCAACCGCACCCCACAGCTATTTCATGTGTGAAG	720
Db	2760	ACATGTGTGCGCCCGGGGGCGGCCAACCGCACCCCACAGCTATTTCATGTGTGAAG	2819
Qy	721	TCCAATAFACAGTCTCTTAATGTGTGAACACTGTCACATCCGGGAGTTTCCACCGGGTCAAG	780
Db	2820	TCCAATAFACAGTCTCTTAATGTGTGAACACTGTCACATCCGGGAGTTTCCACCGGGTCAAG	2879
Qy	781	TGGGAGACATCGGCACTGGGATCAGCAGCCAGATCCGAGCTGGAGCTTCAGCTTGGTCA	840
Db	2880	TGGGAGACATCGGCACTGGGATCAGCAGCCAGATCCGAGCTGGAGCTTCAGCTTGGTCA	2933
Qy	841	CCAAAGCAGGCGACCTGTTTCGTACGATGAGGTGCCAGACTCGGGCATTCGACCTGC	900
Db	2940	CCAAAGCAGGCGACCTGTTTCGTACGATGAGGTGCCAGACTCGGGCATTCGACCTGC	2999
Qy	901	AGTGCACTTGGGCCCTGTATGGCAGCTTCGCTTGSAGCTGGAGGTCAAGCATGGCCAGC	960
Db	3000	AGTGCACTTGGGCCCTGTATGGCAGCTTCGCTTGSAGCTGGAGGTCAAGCATGGCCAGC	3059
Qy	961	TGGAGAACAGGCGCTTAACCTTGCCCTCCACCGCCGGCTCCACTCCCGGAAAGCAGCT	1020
Db	3060	TGGAGAACAGGCGCTTAACCTTGCCCTCCACCGCCGGCTCCACTCCCGGAAAGCAGCT	3119
Qy	1021	TCCTGCTCGGTGCACAGTGTGCCCTGAAAACACAGSCTCAGCCGTTCCAGGGATYTG	1080
Db	3120	TCCTGCTCGGTGCACAGTGTGCCCTGAAAACACAGSCTCAGCCGTTCCAGGGAT--	3177
Qy	1081	NCCAGCCCCCGGCTCARCAGNTGGGAACAGAGGCCCTCGNCAGCNAGNAAGTNGGGG	1140
Db	3178	GCCAGCCCCCGGCTCA--CAGTGGGAACAGGGCCTCG----CAGCAGCAAGTGGGG	3230
Qy	1141	CAAGCNAGANTGCCCTCCAGATTTGCANCTTAGGCCNWTGCCCANCCCTGCTGAADA	1200
Db	3231	GCAAGCAGAATGCCCTCCAGGATTTACACTGAGCCCTGCCCCAC-----CCTCTGA	3284
Qy	1201	AAACAYTNCGCCACAGTGAAGAGACAAAGAGATGNCAGGAGTTNNACCTYGGGGAA	1260
Db	3285	AAAAACATCCGCCACGTGAAGAGACAGAGGATGGC---AGGAGTTACCTGGGGAA	3340
Qy	1261	ACAAAACAGGGATCTTTNTTCTGCCCTGTCTCCAGTNCGAGTTGGGCTGNACCCGTGG	1320
Db	3341	ACAAAACAGGGATCTTT-TTCTGCCCTGTCTCCAGT-CGAGTTTGGGCTGA-CCCCGTG	3396
Qy	1321	ANTCAGTGAACATTTGTTGGCAGANCAGGGGAGAGAGCTTCAGCTGGGTGAGAGGG	1380
Db	3397	GATCAGTGACCATTTGTTGGCAGA-CAGGGGAGAGAGCTTCAGCTGGGTGAGAGGG	3455
Qy	1381	GTGGGCGAGCCCTTCGGGCCCTCACCTNCCAGGCTGCTGTGNAGAGTGTCAAGTGTGA	1440
Db	3456	GTGGGCGAGCCCTTCGGGCCCTCACCTNCCAGGCTGCTGTGNAGAGTGTCAAGTGTGA	3513
Qy	1441	AGGNCACCAANCTCAGNTTTCAGTCAGAACCCAGGTNCAGCAGGTATGCCCGCCGNTA	1500
Db	3514	AGGNCACCAATC--AGGTTTCAGTCAGAACCCAGGT-CAGCAGGTATGCCCGCCGTAG	3569
Qy	1501	GGTTAANNNGGGGCCCTCTNAAAACCCCTTGCTNGGCTNCCACTNGGCGAGCTCANCC	1560
Db	3570	GTTA-----GGGGGCCCTCTAAAACCCCTTGCTGGCTCA---CCTGGCAGCTCA-CCC	3621
Qy	1561	CTTTTTGGGTGTAGGGGAAAAAATGCCCTGACCTCGGGAAGGCTWCCTTGGTAGAATAC	1620
Db	3622	CTTTTTGGGTGTAGGGGAAAAAAGATGCCCTGAACCTTGGGAGGCT-CCCTGGTAGAATAC	3680
Qy	1621	CACACTTTTCAGGTTGTGTGCAACACAGGTCCTGAGTTGACCTCTGTGTTACGCAAGGACC	1680

[illegible]

RESULT 6	ACCESSION
BC035576	VERSION
LOCUS	KEYWORDS
DEFINITION	SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS
TITLE
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 9038 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
Location/Qualifiers
source
1. 4584
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 67.7%; Score 1781.8; DB 6; Length 4584;
Best Local Similarity 89.3%; Pred. No. 0;
Matches 2296; Conservative 5; Mismatches 184; Indels 87; Gaps 35;
Qy 1 CCCCTCTCACAGCCAGGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
Db 2100 CCCCTCTCACAGCCAGGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 2159
Qy 61 CTGAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120
Db 2160 CTGAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 2219
Qy 121 GCCCTTGGAGGGGAGATATAAGAAACCAAGACATCCACGCCAAATCAAGCCAAATTACC 180
Db 2220 GCCCTTGGAGGGGAGATATAAGAAACCAAGACATCCACGCCAAATCAAGCCAAATTACC 2279
Qy 181 ACCAGACCTCTCATGCCAGCCGAGAGAGCTTTTCGCCAAAGGGCCCGAGGGCCCGGCCAG 240
Db 2280 ACCAGACCTCTCATGCCAGCCGAGAGAGCTTTTCGCCAAAGGGCCCGAGGGCCCGGCCAG 2339
Qy 241 CTGAGGAGACAAGGAGAGAGCCCTTAAGTCTCAGCTCTCTCCCAACAGAGCCCCCAG 300
Db 2340 CTGAGGAGACAAGGAGAGAGCCCTTAAGTCTCAGCTCTCTCCCAACAGAGCCCCCAG 2399
Qy 301 AGCCAAACAGTCTCCCTCCCTTGACTTTGACAGAGGAGGAGTCTGGATGTGGGAACCCCT 360
Db 2400 AGCCAAACAGTCTCCCTCCCTTGACTTTGACAGAGGAGGAGTCTGGATGTGGGAACCCCT 2459
Qy 361 TACCTCTGTCTCCTCGAGCCAGCCCTCTGCAGAAACCCAGCTCACCAGAGCGGAAG 420
Db 2460 TACCTCTGTCTCCTCGAGCCAGCCCTCTGCAGAAACCCAGCTCACCAGAGCGGAAG 2519
Qy 421 CAACCGTCCCGAGCAGGAATCGCAGCAGCTGGAATAGAAATTATTCCTCAACAGCCCTGT 480
Db 2520 CAACCGTCCCGAGCAGGAATCGCAGCAGCTGGAATAGAAATTATTCCTCAACAGCCCTGT 2579
Qy 481 CCAGACCATTTCTCTGAGAGAGCAGAGCAAAATCTCTGTGCTCAGATCGACAGCC 540
Db 2580 CCAGACCATTTCTCTGAGAGAGCAGAGCAAAATCTCTGTGCTCAGATCGACAGCC 2639
Qy 541 TCTCCCTGTGGATGACAGTGAGAGAACCCATCAAGGGCTCTCAAGCTCGCGGACA 600
Db 2640 TCTCCCTGTGGATGACAGTGAGAGAACCCATCAAGGGCTCTCAAGCTCGCGGACA 2699
Qy 601 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTCCAGCTGGA 660
Db 2700 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTCCAGCTGGA 2759
Qy 661 ACATGGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 720
Db 2760 ACATGGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 2819
Qy 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCGGGAGTTCCACGGGTCAAAG 780
Db 2820 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCGGGAGTTCCACGGGTCAAAG 2879
Qy 781 TGGAGACATCGCCACTGGGCATCAGCAGCCAGATCCCGAGCTGCAGCCCTTCAGCTTGGTCA 840

Db 2880 TGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCGAGCTTCAGCTTGGTCA 2939
Qy 841 CCAAAGACGGGAGCCCTGTTTCGCTTACACATGAGAGTGGCAGATCGGGCATCCACCTGC 900
Db 2940 CCAAAGACGGGAGCCCTGTTTCGCTTACACATGAGAGTGGCAGATCGGGCATCCACCTGC 2999
Qy 901 AGTGACACTGGCCCTTGATGGCAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 960
Db 3000 AGTGACACTGGCCCTTGATGGCAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 3059
Qy 961 TGGAGAACAGCCCTTAACCTCGCCCTCCACGGCCGGCTCCACACTGCGGAAAGCAGCCT 1020
Db 3060 TGGAGAACAGCCCTTAACCTCGCCCTCCACGGCCGGCTCCACACTGCGG--AAGCAGCCT 3118
Qy 1021 TCTGTCTCGGTGACAGCATGCTGCCCTGAAAACACAGAGCTCAGCCGTTTCCAGGGGATVTG 1080
Db 3119 TCTGTCTCGGTGACAGCATGCTGCCCTG--AAACACAGGCTCAGCCGTTTCCAGGG--ATC 3174
Qy 1081 NCCAGCCCCCGGCTCARGAGTGGGAAACAGGCGCTTCGNAGCAGNAGCNAGGTNGGGG 1140
Db 3175 TGCAGCCCCCGGCTCAGCAGTGGGACCCAGGGCTTCGACAGCAAG-----GTGGG 3227
Qy 1141 CAAGCNAGATGCTCTCCAGGATTTTCANCTGAGCCNTGCCCCANCCCTGCTGAADA 1200
Db 3228 GCAAGCAGAAATGCTCTCCAGGATTTTCACAC--TGAGCCCTGCCCCACCTGCTGTGAAA 3284
Qy 1201 AAACAYTNCGCCACCTGAGAGACAGAGGAGTGCNAGGAGTTNNACCTTGGGAA 1260
Db 3285 AACACT--CCGCCACGTGAAGAGACA---GAGGAGTGGCAGGAGTTACTTCGGGAAA 3338
Qy 1261 ACAAACAGGGATCTTTNTTCTGCCCCCTGCTCCAGTNCAGTTCGCTGACCCGCTTGG 1320
Db 3339 CAACAGGATCTTCTGCCCCCTGCTCCAG-----TCGAGTTGCGCTGACCCGCTTGG- 3390
Qy 1321 ANTCACTGACATTTGTTGGCAGANCAAGGGAGAGAGCTTTCAGCTTGGGTCAAGAGGG 1380
Db 3391 -ATCAGTGAACATTTGCTGGCAGA-CAGGGGAGAGAGCTTCCAGCTTGGGTCAAGAGGG 3448
Qy 1381 GTGGGAGAGCCCTTCGGCCCTCACCCCTNCCAGCTCTGTGNAGAGTGTCAAGTGTGA 1440
Db 3449 GTGGGAGAGCCCTTCGGCCCTCACCCCT--CCAGGCTGCTGTG--AGAGTGTCAAGTGTGA 3506
Qy 1441 AGGNCCTCAAACTCAGGNTTCAGTGCAGAACACAGGTNCAAGCAGTATGCCCCCGGNTA 1500
Db 3507 AGGCCCAAACTC---AGGTTCACTGACAGAACACAGGT--CAGCAGTATGCCCCCGCTAG 3562
Qy 1501 GGTAAANNNGGGGCGCTCTNAAACCCCTTCCTNNGGCTNCACTNNGGCCAGCTCAACCC 1560
Db 3563 GTTAA-----GGGGCGCCCTCTAAACCCCTTGCCTGGCCCTCA---CCTGGCCAGCTCA-CCC 3614
Qy 1561 CTTTGGGTGTAGGGGAAAAGATGCTGACCTTGGGAAGGCTWCCCTGCTAGATATACAC 1620
Db 3615 CTTTGGGTGTAGGGGAAAAGATGCTGACCTTGGGAAGGCT--CCCTGCTAGATATACAC 3673
Qy 1621 CACACTTTTCAGGTGTGTCACACAGGTCTCTGAGTTGACCTCTGTTTTCAGCAGAGACC 1680
Db 3674 CACACTTTTCAGGTGTGTCACACAGGTCTCTGAGTTGACCTCTGTTTTCAGCAGAGACC 3733
Qy 1681 AAAGAAGGTGTGAAGTGAAGTGTTCATGTCNCCCAGACATGTCGCCCTTTGCTGCTG 1740
Db 3734 AAAGAAGGTGTGAAGTGAAGTGTTCATGTCNCCCAGACATGTCGCCCTTTGCTGCTG 3792
Qy 1741 GCTACACATCTTCCAGAGCAGAGCCCGGAGCCCTTCAGGCCAGCAGCTGCCCCAG 1800
Db 3793 GCTACACATCTTCCAGAGCAGAGCCCGGAGCCCTTCAGGCCAGCAGCTGCCCCAG 3852
Qy 1801 ACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAGAGTGAAGGTGATGACAGATATGCTGTA 1860
Db 3853 ACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAGAGTGAAGGTGATGACAGATATGCTGTA 3912
Qy 1861 CAGGAACAGTCTCTGGATGGACATGATGATGCTGCTNAAAGNAAAGCAGCAGAGAGAGCT 1920


```
/clone_lib="TBS02"
/notes="cloning vector: pME18SFL3"
218..1651
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAD18584.1"
/db_xref="GI:47077392"
/translation="MKLIRSSSCHVGSARKLHWRAGOTEPGWAGTRHDSVPVLPPIRG
GLADNVLLSDGSHALCDRGHAVCLQPDGLGKSLITGDDYIPGTEPHMAPEVVLGRS
CDKAVDWSSCCMLHMLUNGHPWTQFFRGGLKSLIAEPPEPPREIPPSCAPLTAQAI
OQELRKEPIHRSAAELGKVNRLAQVGLKSPMRGEYKPRHPHPNPQANHOTLHA
QPRELPAPRGPRAEETGRAPKLQPLPPEPPKPSPLTLISKEESNMWELPLIS
SLEPAPANPSRPERKATVPEQLQLEIELFLNSLSQPFSLERQEQILSLISLISLS
LSDSEKNPSKSSRDTLSSGVHSSWSSQAEARSSNMWVLARGRPDTPSYFNGVK
VQIQSLNGEHLHIREPHRVKVGDIATGIISSQIPAAAFLSVTKDQPVRYDNEVPDSGI
DLQCTLAPDGSFAWSVRVKGQLENRP"

ORIGIN
Query Match 67.3%; Score 1771.6; DB 9; Length 3152;
Best Local Similarity 89.2%; Pred No. 0;
Matches 2287; Conservative 5; Mismatches 186; Indels 87; Gaps 35;

QY 1 CCCCTCTCACAGCCAGGCGCATCCAGAGGGGCTCAGGAAAGAGCCCATCCACCGGTGT 60
DB 675 CCCCCTCTCACAGCCAGGCGCATCCAGAGGGGCTCAGGAAAGAGCCCATCCACCGGTGT 734

QY 61 CTGCAGCGAGCTGGAGGAGGAGTGAACCGGCACTACGCAAGTGGGAGGTCTGAAGA 120
DB 735 CTGCAGCGAGCTGGAGGAGGAGTGAACCGGCACTACGCAAGTGGGAGGTCTGAAGA 794

QY 121 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAATCAAGCCAAATTACC 180
DB 795 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAATCAAGCCAAATTACC 854

QY 181 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTTCGCAAGGGCCCGAGGCCCGGCGAG 240
DB 855 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTTCGCAAGGGCCCGAGGCCCGGCGAG 914

QY 241 CTGGAGAGACACAGGAGAGCCCTTAAGCTCCAGCCCTCTCTCCACAGAGCCCGCCAG 300
DB 915 CTGGAGAGACACAGGAGAGCCCTTAAGCTCCAGCCCTCTCTCCACAGAGCCCGCCAG 974

QY 301 AGCCAAACAAAGTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCT 360
DB 975 AGCCAAACAAAGTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCT 1034

QY 361 TACCTCTGTCTCTCTGTGAGAGCAGCCCTTGCCAGAAACCCCAAGCTTCAAGAGCGGAAAG 420
DB 1035 TACCTCTGTCTCTCTGTGAGAGCAGCCCTTGCCAGAAACCCCAAGCTTCAAGAGCGGAAAG 1094

QY 421 CAACCGTCCGGAGCAGGAACTGCAGCAGCTGGAATAGAAATATTCTCTCAACAGCCTGT 480
DB 1095 CAACCGTCCGGAGCAGGAACTGCAGCAGCTGGAATAGAAATATTCTCTCAACAGCCTGT 1154

QY 481 CCCAGCAATTTCTCTGAGGAGCAGGAGCAAAATCTCTCGTGCCTCAGCATCGACAGCC 540
DB 1155 CCCAGCAATTTCTCTGAGGAGCAGGAGCAAAATCTCTCGTGCCTCAGCATCGACAGCC 1214

QY 541 TCTCCCTGTGCGATGACAGTGAGAGAACCCATCAAGAGCCTCTCAAGCTCGCGGAGCA 600
DB 1215 TCTCCCTGTGCGATGACAGTGAGAGAACCCATCAAGAGCCTCTCAAGCTCGCGGAGCA 1274

QY 601 CCTGAGCTCAGCGTACACTCTCTGAGCAGCAGCCGAGGCTCGAAGCTCCAGCTGGA 660
DB 1275 CCTGAGCTCAGCGTACACTCTCTGAGCAGCAGCCGAGGCTCGAAGCTCCAGCTGGA 1334

QY 661 ACATGTGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTATTTCAATGGTGTGAAG 720
DB 1335 ACATGTGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTATTTCAATGGTGTGAAG 1394

QY 721 TCCAAATACAGTCTCTTAATGTGGAACACCTGCACATCCGGGAGTTCCACCGGGTCAAAG 780
DB 1395 TCCAAATACAGTCTCTTAATGTGGAACACCTGCACATCCGGGAGTTCCACCGGGTCAAAG 1454
```

```
QY 781 TGGGAGACATCGCCACCTGGCATCAGCAGCAGATCCAGAGCTTCCAGCTTGCAGCTTGGTCA 840
DB 1455 TGGGAGACATCGCCACCTGGCATCAGCAGCAGATCCAGAGCTTGCAGCTTGGTCA 1514

QY 841 CCAAAGACGGGACGCTGTTGCTGTACACATGAGAGTGCACAGCTCGGGCATACACCTGC 900
DB 1515 CCAAAGACGGGACGCTGTTGCTGTACACATGAGAGTGCACAGCTCGGGCATACACCTGC 1574

QY 901 AGTCACACTCGGCCCTGATGAGGAGCTTCGCTCGAGCTGGAGGTCAAGCATGCGCAGC 960
DB 1575 AGTCACACTCGGCCCTGATGAGGAGCTTCGCTCGAGCTGGAGGTCAAGCATGCGCAGC 1634

QY 961 TGGAGAACAGGCGCCCTAACCCCTGCCCTCCACCGCGGCTCCACACTGCCG -AAGCAGCCT 1020
DB 1635 TGGAGAACAGGCGCCCTAACCCCTGCCCTCCACCGCGGCTCCACACTGCCG -AAGCAGCCT 1693

QY 1021 TCCTGTCTGTGTCACGATGTCGCTCGAATAACAAGCTCAGCCGTTCCTCCAGGGGATYTG 1080
DB 1694 TCCTGTCTGTGTCACGATGTCGCTCGAATAACAAGCTCAGCCGTTCCTCCAGGG -ATC 1749

QY 1081 NCCAGCCCCCGGCTCARGCAGTGGGAACAGGGGCTCCGACGACGACGAGGTTGGGG 1140
DB 1750 TGCAGGCCCCCGGCTCARGCAGTGGGAACAGGGGCTTCGACGACGAG -GTGGGG 1802

QY 1141 CAAGCAGAAATGCTCTCCAGGATTTACANCTTGAGCCCTGCCCCANCCCTGCTGAADA 1200
DB 1803 GCAAGCAGAAATGCTCTCCAGGATTTACACC -TGAGCCCTGCCCCACCTGCTGAGAA 1859

QY 1201 AAACAYNVCCGACGTCGTGAAGAGACAGAGAGAGATGNCAGGATTTNNACTTGGGAA 1260
DB 1860 AACACT -CCGCCACGTGAAGAGACA -GAGGAGGATGCGCAGGAGTTACTCGGGAAA 1913

QY 1261 ACAAAACAGGATCTTTNTTCTGCCCTGCTCCAGTNCAGTTCGCCCTGNAACCCGCTTGG 1320
DB 1914 CAACAGAGATCTTCTGCCCCCTGCTCCAG -TCAGTTCGCTTGA -CCGCTT -G 1964

QY 1321 ANTCACTGACCATTTGTTGGCAGANCAGGGGAGAGACGCTTCCAGCTTGGGTTCAGAAAGG 1380
DB 1965 GATCAGTGACCATTTGCTGGCAGA -CAGGGAGAGCAGCTTCCAGCTTGGGTTCAGAAAGG 2023

QY 1381 GTGGGCGAGCCCTTCGGCCCTCACCCTNCAGAGTCTGTGNAGAGTCAAGTGTCTGA 1440
DB 2024 GTGGGCGAGCCCTTCGGCCCTCACCCT -CCAGGCTGCTGTG -AGAGTCAAGTGTCTGA 2081

QY 1441 AGGNCNCCAAANCTCAGNTTTCAGTGCAGAACACAGTNCAGAGTATGCCGCCGNTA 1500
DB 2082 AGGCCCCAAACTC -AGGTTTCAGTGCAGAACACAGT -CAGCAGTATGCCGCCGCTAG 2137

QY 1501 GGTAAANNNGGGGGCCCTCTNAAACCCCTTGCTNGCCCTNCACCTNGGCCAGCTCANCCC 1560
DB 2138 GTTAA -GGGGGGCCCTTAAACCCCTTGCTGGCCCTCA -CCTGGCCAGCTCA -CCC 2189

QY 1561 CTTTGGGTGTAGGGGAAAAGAAATGCCCTGACCTCGGGAAGGCTWCCCTGGTAGAATACAC 1620
DB 2190 CTTTGGGTGTAGGGGAAAAGAAATGCCCTGACCTCGGGAAGGCT -CCCTGGTAGAATACAC 2248

QY 1621 CACACTTTTCAGTTGTTGCAACACAGTCTGAGTTCACCTCTGCTTTCAGCAGAGGACC 1680
DB 2249 CACACTTTTCAGTTGTTGCAACACAGTCTGAGTTCACCTCTGCTTTCAGCAGAGGACC 2308

QY 1681 AAAAGAGGTGTAAAGTGAAGTGGTTCTCAGTNCCTCCAGACATGTCGCCCTTTGCTGTG 1740
DB 2309 AAAAGAGGTGTAAAGTGAAGTGGTTCTCAGT -CCCCAGACATGTCGCCCTTTGCTGTG 2367

QY 1741 GCTACACTCTTCCCCAGAGCAGGAGCCCGAGCCCTTTCAGGCCCGCAGCACTGCCCCAG 1800
DB 2368 GCTACACTCTTCCCCAGAGCAGGAGCCCGAGCCCTTTCAGGCCCGCAGCACTGCCCCAG 2427

QY 1801 ACTCCTGCGCACTCAGTTCCTCTCATCTGTAAAGGTGAAGGGTGTATGCGAGGATATGCTGA 1860
DB 2428 ACTCCTGCGCACTCAGTTCCTCTCATCTGTAAAGGTGAAGGGTGTATGCGAGGATATGCTGA 2487
```



```

Db 2946 GCCAGCCCCCGGCTCA--CAGTGGAAACAGGGGCTCG-----CAGCAACAAGTGGG 2998
QY 1141 CAAGCNAGAATGCTCCAGGATTTACANCTGAGCCNTGCCCCANCCCTGCTGAADA 1200
Db 2999 GCAAGCAGAATGCTCCAGGATTTTACACCTGAGCCCTGCCCCAC-----CCTGCTGA 3052
QY 1201 AAACAATNCCGCGCAGTGAAGAGACAGAAAGAGGATGNCAGAGGATTTNNACTTGGGAA 1260
Db 3053 AAAACATCCCGCAGCTGAAGAGACAGAAAGAGGATGSC----AGGAGTTACCTGGGAA 3108
QY 1261 ACAAAACAGGGATCTTTTCTGCCCCCTGCTCCAGTNCAGTGGCCTG 1309
Db 3109 ACAAAACAGGGATCTTTTCTGCCCCCTGCTCCAGT-CGAGTTGGCCTG 3155

RESULT 11
AR063065
LOCUS AR063065 3156 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5844073.
ACCESSION AR063065
VERSION AR063065.1 GI:5990756
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS Rothe,M. and Wu,L.
TITLE Human NIK proteins
JOURNAL Patent: US 5844073-A 1 01-DEC-1998;
FEATURES
    Location/Qualifiers
        1..3156
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
Query Match 43.1%; Score 1133.4; DB 6; Length 3156;
Best Local Similarity 94.8%; Pred. No. 1.2e-279;
Matches 1241; Conservative 4; Mismatches 43; Indels 21; Gaps 7;

QY 1 CCCCTCTACAGCCAGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCCGCTGT 60
Db 1868 CCCCTCTACAGCCAGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCCGCTGT 1927
QY 61 CTGAGCGAGCTGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120
Db 1928 CTGAGCGAGCTGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 1987
QY 121 GCCCTTGGAGGGGAGAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC 180
Db 1988 GCCCTTGGAGGGGAGAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC 2047
QY 181 ACCAGACCTTCATGCCAGCGAGAGAGCTTTGCGCAAGGCCCCAGGGCCCGGCCAG 240
Db 2048 ACCAGACCTTCATGCCAGCGAGAGAGCTTTGCGCAAGGCCCCAGGGCCCGGCCAG 2107
QY 241 CTGAGGAGACAACAGCAGAGCCCTTAAGCTCAGCTCCTCTCCACACAGAGCCCCAG 300
Db 2108 CTGAGGAGACAACAGCAGAGCCCTTAAGCTCAGCTCCTCTCCACACAGAGCCCCAG 2167
QY 301 AGCCAAACAAAGTCTCCTCCCTTGACTTTGAGAAGGAGGAGTCTGGGATGGGAAACCT 360
Db 2168 AGCCAAACAAAGTCTCCTCCCTTGACTTTGAGAAGGAGGAGTCTGGGATGGGAAACCT 2227
QY 361 TACCTCTGTCTCTCTGGAGCAGGCCCTGCGCAGAAACCCAGCTCACCAGAGCGGAAAG 420
Db 2228 TACCTCTGTCTCTCTGGAGCAGGCCCTGCGCAGAAACCCAGCTCACCAGAGCGGAAAG 2287
QY 421 CAACCGTCCCGAGCAGGAACTGCGAGCAGCTGGAATAGAAATTATTCTCAACAGCCTGT 480
Db 2288 CAACCGTCCCGAGCAGGAACTGCGAGCAGCTGGAATAGAAATTATTCTCAACAGCCTGT 2347
QY 481 CCCAGCCATTTTCTCTGGAGGAGCAGGAGCAAAATTCTCTCGTGCCTCAGCATCGACGCC 540
```

```

Db 2348 CCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGCCTCAGCATCGACAGCC 2407
QY 541 TCTCCCTGTGCGATGACAGTGAAGAACCCATCAAAAGGCTCTCAAAAGCTCGCGGGACA 600
Db 2408 TCTCCCTGTGCGATGACAGTGAAGAACCCATCAAAAGGCTCTCAAAAGCTCGCGGGACA 2467
QY 601 CCCTGAGCTCAGGCGGTACACTCTCTGAGCAGCAGGCGCGAGGCTCGAAGCTCCAGCTGGA 660
Db 2468 CCCTGAGCTCAGGCGGTACACTCTCTGAGCAGCAGGCGCGAGGCTCGAAGCTCCAGCTGGA 2527
QY 661 ACATGGTGTGGCCCGGGGGCGGCCACCCGACACCCCAAGCTATTTCATGGTGTGAAG 720
Db 2528 ACATGGTGTGGCCCGGGGGCGGCCACCCGACACCCCAAGCTATTTCATGGTGTGAAG 2587
QY 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCCACCGGGTCAAG 780
Db 2588 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCCACCGGGTCAAG 2647
QY 781 TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCCGAGCTGCAGCCTTCAGCTTGTCA 840
Db 2648 TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCCGAGCTGCAGCCTTCAGCTTGTCA 2707
QY 841 CCAAAGACGGGCGAGCTGTTTCGCTACGACATGGAAGTCCAGACTCGGGCATCGACCTGC 900
Db 2708 CCAAAGACGGGCGAGCTGTTTCGCTACGACATGGAAGTCCAGACTCGGGCATCGACCTGC 2767
QY 901 AGTGCACACTGGCCCTTGATGGCAGCTTTCGCTGGAGCTGAGGCTCAAGCATGGCCAGC 960
Db 2768 AGTGCACACTGGCCCTTGATGGCAGCTTTCGCTGGAGCTGAGGCTCAAGCATGGCCAGC 2827
QY 961 TGGGAAACAGGCCCTTAACCTGCTCCATCCACCGCGGCTCCACACTGCCGGAAGCAGCCT 1020
Db 2828 TGGGAAACAGGCCCTTAACCTGCTCCATCCACCGCGGCTCCACACTGCCGGAAGCAGCCT 2887
QY 1021 TCCTGCTCGTGCAGATGTCGCCCTGAAACACAGAGCTCAGCGCTTCCAGGGGATTTG 1080
Db 2888 TCCTGCTCGTGCAGATGTCGCCCTGAAACACAGAGCTCAGCGCTTCCAGGGGATTT-- 2945
QY 1081 NCCAGCCCCCGGCTCARGAGTGGGAACAGGCGCTCGNCGAGNAGNAGGTTNGGGG 1140
Db 2946 GCCAGCCCCCGGCTCA--CAGTGGAAACAGAGGCTCG-----CCTGCTGA 2998
QY 1141 CAAGCAGAAATGCTCTCCAGGATTTTCAACNCTGAGCCNTGCCCANCCCTGCTGAADA 1200
Db 2999 GCAAGCAGAAATGCTCTCCAGGATTTTCAACCTGAGCCCTGCCCCAC-----CCTGCTGA 3052
QY 1201 AAACAATNCCGCGCAGTGAAGAGACAGAGAGGATGNCAGAGGATTTNNACTTGGGAA 1260
Db 3053 AAAACATCCCGCAGCTGAAGAGACAGAAAGAGGATGSC----AGGAGTTACCTGGGAA 3108
QY 1261 ACAAAACAGGGATCTTTTCTGCCCCCTGCTCCAGTNCAGTGGCCTG 1309
Db 3109 ACAAAACAGGGATCTTTTCTGCCCCCTGCTCCAGT-CGAGTTGGCCTG 3155

RESULT 12
AR068515
LOCUS AR068515 3156 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5854003.
ACCESSION AR068515
VERSION AR068515.1 GI:6000722
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS Rothe,M. and Wu,L.
TITLE Screening method for agents that modulate human NIK activity
JOURNAL Patent: US 5854003-A 1 29-DEC-1998;
FEATURES
    Location/Qualifiers
        1..3156
            /organism="unknown"
            source
```

```
/mol type="unassigned DNA"
```

ORIGIN

Query Match 43.1%; Score 1133.4; DB 6; Length 3156;
Best Local Similarity 94.8%; Pred. NO. 1.2e-279;
Matches 1241; Conservative 4; Mismatches 43; Indels 21;

Qy	1	CCCTCTCACAGCCCGAGCCCATCAAGAGGGGCTGAGAAAGACCCCATCAACCGCGTGT	60
Db	1868	CCCTCTCACAGCCCGAGCCATCAAGAGGGGCTGAGAAAGACCCCATCAACCGCGTGT	1927
Qy	61	CTGAGCGGAGCTGGAGGGAAGTGAACCGGGCACTACAGCAAGTCGAGGTCGGAAGA	120
Db	1928	CTGAGCGGAGCTGGAGGGAAGTGAACCGGGCACTACAGCAAGTCGAGGTCGGAAGA	1987
Qy	121	GCCCTTGAGGGGGAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAATTACC	180
Db	1988	GCCCTTGAGGGGGAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAATTACC	2047
Qy	181	ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCGAGGCCCGGCCAG	240
Db	2048	ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCGAGGCCCGGCCAG	2107
Qy	241	CTGAGGAGACAACAGGCAGAGCCCTTAAGCTTCAGCCTCTCTCCACACAGAGCCCCCAG	300
Db	2108	CTGAGGAGACAACAGGCAGAGCCCTTAAGCTTCAGCCTCTCTCCACACAGAGCCCCCAG	2167
Qy	301	AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCCCT	360
Db	2168	AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCCCT	2227
Qy	361	TACTCTGTCTCTCCCTGGAGCGACCCCTGCAGAAACCCAGCTCACACAGAGCGGAAG	420
Db	2228	TACTCTGTCTCTCCCTGGAGCGACCCCTGCAGAAACCCAGCTCACACAGAGCGGAAG	2287
Qy	421	CAACCGTCCGGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATTAATCTCAACAGCCTGT	480
Db	2288	CAACCGTCCGGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATTAATCTCAACAGCCTGT	2347
Qy	481	CCAGCCATTTTCTCTGGAGGACAGAGCAAAATTCCTCTGCCTCAGCATCGACAGCC	540
Db	2348	CCAGCCATTTTCTCTGGAGGACAGAGCAAAATTCCTCTGCCTCAGCATCGACAGCC	2407
Qy	541	TCTCCCTGTCCGATGACAGTGAGAGAACCCATCAAAAGGCTCTCAAAGCTCGCGGACA	600
Db	2408	TCTCCCTGTCCGATGACAGTGAGAGAACCCATCAAAAGGCTCTCAAAGCTCGCGGACA	2467
Qy	601	CCCTGAGCTCAGCGGTACATCTCTGGAGCAGCAGGCCGAGGCTCGAGATCTCAGTGGA	660
Db	2468	CCCTGAGCTCAGCGGTACATCTCTGGAGCAGCAGGCCGAGGCTCGAGATCTCAGTGGA	2527
Qy	661	ACATGTGTCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAG	720
Db	2528	ACATGTGTCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAG	2587
Qy	721	TCCAAATACAGTCTCTTAATGGTGAAACACTGACATCCGGGAGTTCACACGGGTCAAAG	780
Db	2588	TCCAAATACAGTCTCTTAATGGTGAAACACTGACATCCGGGAGTTCACACGGGTCAAAG	2647
Qy	781	TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTGCAGCCTTCAGCTTGGTCA	840
Db	2648	TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTGCAGCCTTCAGCTTGGTCA	2707
Qy	841	CCAAAGACGGCAGCCTGTTGCTTACGACATGAGAGTGCAGACTCGGGCATCGACTGC	900
Db	2708	CCAAAGACGGCAGCCTGTTGCTTACGACATGAGAGTGCAGACTCGGGCATCGACTGC	2767
Qy	901	AGTGCACTACCTGGCCCTGTATGGAGGCTTCGCTTGGAGCTGGAGGGTCAAGCATGGCCAGC	960
Db	2768	AGTGCACTACCTGGCCCTGTATGGAGGCTTCGCTTGGAGCTGGAGGGTCAAGCATGGCCAGC	2827
Qy	961	TGGAGAACAGGCCCTTAACCCCTGCCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT	1020

D	b		2828	TGAGAAACAGGCCCTTAACCCCTGCCTTCCACCGCGGGCTCCACACTGCGGAAAAGCAGCCT	2888
Q	y		1021	TCCTGCTCGGTGACAGATGCTGCCTTGAAAAACACAGCTCAGCCGTTTCCCAGGGGATYTG	1080
D	b		2888	TCCTGCTCGGTGACAGATGCTGCCTTGAAAAACACAGCTCAGCCGTTTCCCAGGGGAT--	2945
Q	y		1081	NCCAGCCCCCGGCTCARCAGNTGGGAACACAGGGCCTCGNCAGCNAGCNAAAGTNGGGG	1140
D	b		2946	GCCAGCCCCCGGCTCA--CAGTGGGAACACAGGGCCTCG-----CAGCACAAGTGGGG	2998
Q	y		1141	CAACGNAGAATGCTCCAGAGATTTCACANCCTGAGGCCNWTGCCANCCCTCTCTAADA	1200
D	b		2999	GCAAGCAGAATGCTCCAGAGATTTCACACCCTGAGCCCTGCCCCAC-----CCTGCTGA	3052
Q	y		1201	AAACAYTNC CGCACGTAAGAGACAGAGAGGAGATGGNCAGGAGTTNNACCTYGGGAA	1260
D	b		3053	AAAACATCGCCACGTAAGAGACAGAGAGAGATGGC---AGAGTTACCTGGGGAA	3108
Q	y		1261	ACAAAACAGGATCTTTTNTTCTCCCTGCTCCAGTNCGAGTTGGCCTG	1309
D	b		3109	ACAAAACAGGATCTTTT-TTCTGCCCTGCTCCAGT-CGAGTTGGCCTG	3155
RESULT 13					
BD062401					
LOCUS					
DEFINITION NIK proteins, nucleic acids and methods.					
ACCESSION BD062401					
VERSION BD062401.1 GI:22608004					
KEYWORDS JP 2001510348-A/1.					
SOURCE Homo sapiens (human)					
ORGANISM Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE					
AUTHORS Rothe,M. and Wu,L.					
TITLE NIK proteins, nucleic acids and methods					
JOURNAL Patent: JP 2001510348-A 1 31-JUL-2001;					
TULARIK INC					
COMMENT					
PN JP 2001510348-A/1					
PD 31-JUL-2001					
PF 02-JUL-1998 JP 1999507409					
PR 03-JUL-1997 US 08/887518					
PI MIKE ROTHE, LIN WU					
PC					
C07K7/06,C07K7/08,C07K14/47,C12N1/00,C12N5/10,C12N15/12,C12N15/ PC					
63,					
CC C12P21/02,G01N33/53					
CC Strandedness: Double;					
CC Topology: Linear;					
FH Key Location/Qualifiers.					
FEATURES					
source					
1..3156					
/organism="Homo sapiens"					
/mol_type="genomic DNA"					
/db_xref="taxon:9606"					
ORIGIN					
Query Match 43.1%; Score 1133.4; DB 6; Length 3156;					
Best Local Similarity 94.8%; Pred.No.1.2e-279;					
Matches 1241; Conservative 4; Mismatches 43; Indels 21; Gaps .					
Q	y	1	CCCTCTCAGCCCCCAGGCCATCCAGAGGGGCTGAGGAAGGCCCATCCACCGCGTGT	60	
D	b	1868	CCCTCTCAGCCCCCAGGCCATCCAGAGGGGCTGAGGAAGGCCCATCCACCGCGTGT	1927	
Q	y	61	CTGCAGCGAGCTGGGAGGAAGGTCAAACGGGCCTACAGCAAGTGGGAGGTCTGAAGA	120	
D	b	1928	CTGCAGCGAGCTGGGAGGAAGGTCAAACGGGCCTACAGCAAGTGGGAGGTCTGAAGA	1987	
Q	y	121	GCCTTTGGGGGAGAAATAAAGAACCAAGACATCCACCGCAATCAAGCCAATTACC	180	
D	b	1988	GCCTTTGGGGGAGAAATAAAGAACCAAGACATCCACCGCAATCAAGCCAATTACC	2047	

181 ACCAGACCTCCATGCCCCAGCGAGAGCTTTTCCCAAGGGCCCCCGGCCAG 240
Db
2048 ACCAGACCTCCATGCCAGCGAGAGCTTTTCCCAAGGGCCCCCGGCCAG 2107
Qy
241 CTGAGGAGACACAGGAGAGCCCTTACAGCTCCAGCTCTCCACACAGAGCCCGCAG 300
Db
2108 CTGAGGAGACACAGGAGAGCCCTTACAGCTCTCCACACAGAGCCCGCAG 2167
Qy
301 AGCCAAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGAGTCTGGGATGCGAAACCT 360
Db
2168 AGCCAAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGAGTCTGGGATGCGAAACCT 2227
Qy
361 TACCTCTGTCTCTCTCTGAGCGAGCCCTTCCAGAGAACCCAGCTCCACAGAGCGGAAG 420
Db
2228 TACCTCTGTCTCTCTCTGAGCGAGCCCTTCCAGAGAACCCAGCTCCACAGAGCGGAAG 2287
Qy
421 CAACCGTCCCGAGCAGGAACCTGACAGCAGCTGGAATAAGAAATTAATTCCTCAACAGCCTGT 480
Db
2288 CAACCGTCCCGAGCAGGAACCTGACAGCAGCTGGAATAAGAAATTAATTCCTCAACAGCCTGT 2347
Qy
481 CCCAGCAATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATCGACAGCC 540
Db
2348 CCCAGCAATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATCGACAGCC 2407
Qy
541 TCTCCCTGTGATGACAGTGAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGACA 600
Db
2408 TCTCCCTGTGATGACAGTGAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGACA 2467
Qy
601 CCTGAGCTCAGCGGTACACTCTCTGAGCAGCAGCGCCGAGGCTCGAAGCTCCAGCTGGA 660
Db
2468 CCTGAGCTCAGCGGTACACTCTCTGAGCAGCAGCGCCGAGGCTCGAAGCTCCAGCTGGA 2527
Qy
661 ACATGTGTGTCGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTGTGAAG 720
Db
2528 ACATGTGTGTCGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTGTGAAG 2587
Qy
721 TCCAAATACAGTCTCTTAATGTGAACCTTCACATCCGGAGTTCCACCGGGTCAAAG 780
Db
2588 TCCAAATACAGTCTCTTAATGTGAACCTTCACATCCGGAGTTCCACCGGGTCAAAG 2647
Qy
781 TGGGAGACATCGCCACTTGGCATCAGCAGCAGATCCAGCTCGAGCTTTCAGCTTGGTCA 840
Db
2648 TGGGAGACATCGCCACTTGGCATCAGCAGCAGATCCAGCTCGAGCTTTCAGCTTGGTCA 2707
Qy
841 CAAAAGACGGGAGCCTGTTTCGTACGATGAGATGAGAGTGCAGACTCGGGCATCGACCTGC 900
Db
2708 CAAAAGACGGGAGCCTGTTTCGTACGATGAGATGAGAGTGCAGACTCGGGCATCGACCTGC 2767
Qy
901 AGTGCACTGGCCCTGATGGCAGCTTCGCCCTGGAGCTGGAGGGTCAAGCATGGCCAGC 960
Db
2768 AGTGCACTGGCCCTGATGGCAGCTTCGCCCTGGAGCTGGAGGGTCAAGCATGGCCAGC 2827
Qy
961 TGGAGAACAGGCGCTTAAACCTGCGCTCCACCGCGGCTCCACACTCGCGGAAAGCAGCT 1020
Db
2828 TGGAGAACAGGCGCTTAAACCTGCGCTCCACCGCGGCTCCACACTCGCGGAAAGCAGCT 2887
Qy
1021 TCCTGCTGGTGCAGATGCTGCTTGAACACAGGCTCAGCGTTCCAGGGGATYG 1080
Db
2888 TCCTGCTGGTGCAGATGCTGCTTGAACACAGGCTCAGCGTTCCAGGGGATYG 2945
Qy
1081 NCCAGCCCCCGCTCAGAGTGGGAAACAGGGCTTCGNCAGCAGCAGGAGTNGGGG 1140
Db
2946 GCCAGCCCCCGGCTCA--CAGTGGGAAACAGGGCTCG----CAGCAGAGGTGGG 2998
Qy
1141 CAAGCAGAAATGCTCCAGGATTCACATTCAGCCCTGAGCCCTCCAGCTCTGAADA 1200
Db
2999 GCAAGCAGAAATGCTCCAGGATTCACATTCAGCCCTGAGCCCTCGCCAC-----CCTGCTGA 3052
Qy
1201 AAACAYTNCGCCACGCTGAAGACAGAGGATGNCAGGAGTNNACCTYGGGAA 1260
Db
3053 AAAAACAATCCGCGACGCTGAAGAGACAGAGGATGCG----AGGAGTTACCTGGGAA 3108

1261 ACAAAACAGGATCTTTTCTTCTCCCTCTCCAGTCCAGTTCGAGTTGGCTG 1309
Db
3109 ACAAAACAGGATCTTTTCTTCTCCCTCTCCAGTTCGAGTTGGCTG 3155
RESULT 14
LOCUS AR429673 2844 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 2 from patent US 6645728.
ACCESSION AR429673
VERSION AR429673.1 GI:40189976
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2844)
AUTHORS Greene,W.C.; Lin,X. and Gelezuinas,R.
TITLE Inhibitor of the inflammatory response induced by TNF.alpha. and IL-1
JOURNAL Patent: US 6645728-A 2 11-NOV-2003;
FEATURES
source Location/Qualifiers
1..2844
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 37.1%; Score 977; DB 6; Length 2844;
Best Local Similarity 100.0%; Pred. No. 1.8e-239;
Matches 977; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCCCTCTCACAGCCAGGCCATCCAGAGGGGCTGAGGAAAGAGCCATCCACGCGTGT 60
Db 1868 CCCCTCTCACAGCCAGGCCATCCAGAGGGGCTGAGGAAAGAGCCATCCACGCGTGT 1927
Qy 61 CTGAGCGGAGCTCGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120
Db 1928 CTGAGCGGAGCTCGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 1987
Qy 121 GCCCTTGGAGGGGAGAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 180
Db 1988 GCCCTTGGAGGGGAGAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 2047
Qy 181 ACCAGACCTCCATGCCAGCGGAGAGCTTTCCGCAAGGGCCCGAGGCCCGCCAG 240
Db 2048 ACCAGACCTCCATGCCAGCGGAGAGCTTTCCGCAAGGGCCCGAGGCCCGCCAG 2107
Qy 241 CTGAGGAGACAAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAG 300
Db 2108 CTGAGGAGACAAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAG 2167
Qy 301 AGCCAAAACAAGTCTCTCTCTTGTACTTTGAGAGGAGAGTCTGGGATGCGGACCT 360
Db 2168 AGCCAAAACAAGTCTCTCTCTTGTACTTTGAGAGGAGAGTCTGGGATGCGGACCT 2227
Qy 361 TACCTCTGTCTCTCTGAGCGAGCCCTTCCAGAGAACCCAGCTCCACAGAGCGGAAG 420
Db 2228 TACCTCTGTCTCTCTGAGCGAGCCCTTCCAGAGAACCCAGCTCCACAGAGCGGAAG 2287
Qy 421 CAACCGTCCCGAGCAGGAACCTGACAGCAGCTGGAATAAGAAATTAATTCCTCAACAGCCTGT 480
Db 2288 CAACCGTCCCGAGCAGGAACCTGACAGCAGCTGGAATAAGAAATTAATTCCTCAACAGCCTGT 2347
Qy 481 CCCAGCAATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATCGACAGCC 540
Db 2348 CCCAGCAATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATCGACAGCC 2407
Qy 541 TCTCCCTGTGATGACAGTGAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGACA 600
Db 2408 TCTCCCTGTGATGACAGTGAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGACA 2467
Qy 601 CCTGAGCTCAGCGGTACACTCTCTGAGCAGCAGCGCCGAGGCTCGAAGCTCCAGCTGGA 660
Db 2468 CCTGAGCTCAGCGGTACACTCTCTGAGCAGCAGCGCCGAGGCTCGAAGCTCCAGCTGGA 2527

repeat_region	complement (14740..14818)	repeat_region	1051	ACACAGGCTCAGCCGTTCCAGGGGATYTGNCAGAGCCGCCGGCTCCAGNTGGGAACC	1110
repeat_region	/rpt_family="MERSA"	repeat_region	1052		1111
repeat_region	complement (15630..15704)	repeat_region	1053	ACACAGGCTCAGCCGTTCCAGGGG---ATCTGCCAGCCGCCGGCTCAGCAGTGGGACCAG	44989
repeat_region	/rpt_family="CA)n"	repeat_region	1054		1112
repeat_region	16564..16662	repeat_region	1055	1111	1113
repeat_region	/rpt_family="MIR"	repeat_region	1056	1111	1114
repeat_region	complement (16921..17215)	repeat_region	1057	1111	1115
repeat_region	/rpt_family="AluJb"	repeat_region	1058	1111	1116
repeat_region	18177..18205	repeat_region	1059	1111	1117
repeat_region	/rpt_family="CA)n"	repeat_region	1060	1111	1118
repeat_region	complement (19855..19953)	repeat_region	1061	1111	1119
repeat_region	/rpt_family="MIR"	repeat_region	1062	1111	1120
repeat_region	20824..21121	repeat_region	1063	1111	1121
repeat_region	/rpt_family="AluSc"	repeat_region	1064	1111	1122
repeat_region	complement (21866..22038)	repeat_region	1065	1111	1123
repeat_region	/rpt_family="MIR"	repeat_region	1066	1111	1124
repeat_region	complement (23516..23817)	repeat_region	1067	1111	1125
repeat_region	/rpt_family="AluJo"	repeat_region	1068	1111	1126
repeat_region	complement (23852..24153)	repeat_region	1069	1111	1127
repeat_region	/rpt_family="AluSx"	repeat_region	1070	1111	1128
repeat_region	complement (25772..25870)	repeat_region	1071	1111	1129
repeat_region	/rpt_family="MER46"	repeat_region	1072	1111	1130
repeat_region	complement (26923..27203)	repeat_region	1073	1111	1131
repeat_region	/rpt_family="AluJb"	repeat_region	1074	1111	1132
repeat_region	27261..27394	repeat_region	1075	1111	1133
repeat_region	/rpt_family="AluJo"	repeat_region	1076	1111	1134
repeat_region	27554..27854	repeat_region	1077	1111	1135
repeat_region	/rpt_family="AluSgI"	repeat_region	1078	1111	1136
repeat_region	27869..28156	repeat_region	1079	1111	1137
repeat_region	/rpt_family="AluSc"	repeat_region	1080	1111	1138
repeat_region	28206..28382	repeat_region	1081	1111	1139
repeat_region	/rpt_family="AluJo"	repeat_region	1082	1111	1140
repeat_region	28587..28882	repeat_region	1083	1111	1141
repeat_region	/rpt_family="AluSx"	repeat_region	1084	1111	1142
repeat_region	28893..29195	repeat_region	1085	1111	1143
repeat_region	/rpt_family="AluSx"	repeat_region	1086	1111	1144
repeat_region	complement (29412..29506)	repeat_region	1087	1111	1145
repeat_region	/rpt_family="LINE2"	repeat_region	1088	1111	1146
repeat_region	29775..30060	repeat_region	1089	1111	1147
repeat_region	/rpt_family="AluY"	repeat_region	1090	1111	1148
repeat_region	30071..30372	repeat_region	1091	1111	1149
repeat_region	/rpt_family="AluSp"	repeat_region	1092	1111	1150
repeat_region	30536..30817	repeat_region	1093	1111	1151
repeat_region	/rpt_family="AluJb"	repeat_region	1094	1111	1152
repeat_region	30859..31080	repeat_region	1095	1111	1153
repeat_region	/rpt_family="MIR"	repeat_region	1096	1111	1154
repeat_region	complement (31106..31144)	repeat_region	1097	1111	1155
repeat_region	/rpt_family="MIR"	repeat_region	1098	1111	1156
repeat_region	32189..32335	repeat_region	1099	1111	1157
repeat_region	/rpt_family="L1LP4"	repeat_region	1100	1111	1158
repeat_region	32370..32670	repeat_region	1101	1111	1159
repeat_region	/rpt_family="AluSg"	repeat_region	1102	1111	1160
repeat_region	32685..32979	repeat_region	1103	1111	1161
repeat_region	/rpt_family="AluSx"	repeat_region	1104	1111	1162
repeat_region	33437..33738	repeat_region	1105	1111	1163
repeat_region	/rpt_family="AluSg"	repeat_region	1106	1111	1164
repeat_region	complement (33995..34315)	repeat_region	1107	1111	1165
repeat_region	/rpt_family="AluSc"	repeat_region	1108	1111	1166
repeat_region	complement (34339..34639)	repeat_region	1109	1111	1167
repeat_region	/rpt_family="AluY"	repeat_region	1110	1111	1168
repeat_region	complement (36539..36839)	repeat_region	1111	1111	1169
repeat_region	/rpt_family="AluSx"	repeat_region	1112	1111	1170
repeat_region	complement (37152..37449)	repeat_region	1113	1111	1171
repeat_region	/rpt_family="AluJo"	repeat_region	1114	1111	1172
repeat_region	37854..38155	repeat_region	1115	1111	1173
repeat_region	/rpt_family="AluSx"	repeat_region	1116	1111	1174
repeat_region	39172..39237	repeat_region	1117	1111	1175
repeat_region	/rpt_family="MIR"	repeat_region	1118	1111	1176
repeat_region	39346..40104	repeat_region	1119	1111	1177
repeat_region	/rpt_family="MER21B"	repeat_region	1120	1111	1178
repeat_region	40186..40475	repeat_region	1121	1111	1179
repeat_region	/rpt_family="AluJo"	repeat_region	1122	1111	1180
repeat_region	40946..41076	repeat_region	1123	1111	1181
repeat_region		repeat_region	1124	1111	1182
		repeat_region	1125	1111	1183
		repeat_region	1126	1111	1184
		repeat_region	1127	1111	1185
		repeat_region	1128	1111	1186
		repeat_region	1129	1111	1187
		repeat_region	1130	1111	1188
		repeat_region	1131	1111	1189
		repeat_region	1132	1111	1190
		repeat_region	1133	1111	1191
		repeat_region	1134	1111	1192
		repeat_region	1135	1111	1193
		repeat_region	1136	1111	1194
		repeat_region	1137	1111	1195
		repeat_region	1138	1111	1196
		repeat_region	1139	1111	1197
		repeat_region	1140	1111	1198
		repeat_region	1141	1111	1199
		repeat_region	1142	1111	1200
		repeat_region	1143	1111	1201
		repeat_region	1144	1111	1202
		repeat_region	1145	1111	1203
		repeat_region	1146	1111	1204
		repeat_region	1147	1111	1205
		repeat_region	1148	1111	1206
		repeat_region	1149	1111	1207
		repeat_region	1150	1111	1208
		repeat_region	1151	1111	1209
		repeat_region	1152	1111	1210
		repeat_region	1153	1111	1211
		repeat_region	1154	1111	1212
		repeat_region	1155	1111	1213
		repeat_region	1156	1111	1214
		repeat_region	1157	1111	1215
		repeat_region	1158	1111	1216
		repeat_region	1159	1111	1217
		repeat_region	1160	1111	1218
		repeat_region	1161	1111	1219
		repeat_region	1162	1111	1220
		repeat_region	1163	1111	1221
		repeat_region	1164	1111	1222
		repeat_region	1165	1111	1223
		repeat_region	1166	1111	1224
		repeat_region	1167	1111	1225
		repeat_region	1168	1111	1226
		repeat_region	1169	1111	1227
		repeat_region	1170	1111	1228
		repeat_region	1171	1111	1229
		repeat_region	1172	1111	1230
		repeat_region	1173	1111	1231
		repeat_region	1174	1111	1232
		repeat_region	1175	1111	1233
		repeat_region	1176	1111	1234
		repeat_region	1177	1111	1235
		repeat_region	1178	1111	1236
		repeat_region	1179	1111	1237
		repeat_region	1180	1111	1238
		repeat_region	1181	1111	1239
		repeat_region	1182	1111	1240
		repeat_region	1183	1111	1241
		repeat_region	1184	1111	1242
		repeat_region	1185	1111	1243
		repeat_region	1186	1111	1244
		repeat_region	1187	1111	1245
		repeat_region	1188	1111	1246
		repeat_region	1189	1111	1247
		repeat_region	1190	1111	1248
		repeat_region	1191	1111	1249
		repeat_region	1192	1111	1250
		repeat_region	1193	1111	1251
		repeat_region	1194	1111	1252
		repeat_region	1195	1111	1253
		repeat_region	1196	1111	1254
		repeat_region	1197	1111	1255
		repeat_region	1198	1111	1256
		repeat_region	1199	1111	1257
		repeat_region	1200	1111	1258
		repeat_region	1201	1111	1259
		repeat_region	1202	1111	1260
		repeat_region	1203	1111	1261
		repeat_region	1204	1111	1262
		repeat_region	1205	1111	1263
		repeat_region	1206	1111	1264
		repeat_region	1207	1111	1265
		repeat_region	1208	1111	1266
		repeat_region	1209	1111	1267
		repeat_region	1210	1111	1268
		repeat_region	1211	1111	1269
		repeat_region	1212	1111	1270
		repeat_region	1213	1111	1271
		repeat_region	1214	1111	1272
		repeat_region	1215	1111	1273
		repeat_region	1216	1111	1274
		repeat_region	1217	1111	1275
		repeat_region	1218	1111	1276
		repeat_region	1219	1111	1277
		repeat_region	1220	1111	1278
		repeat_region	1221	1111	1279
		repeat_region	1222	1111	1280
		repeat_region	1223	1111	1281
		repeat_region	1224	1111	1282
		repeat_region	1225	1111	1283
		repeat_region	1226	1111	1284
		repeat_region	1227	1111	1285
		repeat_region	1228	1111	1286
		repeat_region	1229	1111	1287
		repeat_region	1230	1111	1288
		repeat_region	1231	1111	1289
		repeat_region	1232	1111	1290
		repeat_region	1233	1111	1291
		repeat_region	1234	1111	1292
		repeat_region	1235	1111	1293
		repeat_region	1236	1111	1294
		repeat_region	1237	1111	1295
		repeat_region	1238	1111	1296
		repeat_region	1239	1111	1297
		repeat_region	1240	1111	1298
		repeat_region	1241	1111	1299
		repeat_region	1242	1111	1300
		repeat_region	1243	1111	1301
		repeat_region	1244	1111	1302
		repeat_region	1245	1111	1303
		repeat_region	1246	1111	1304
		repeat_region	1247	1111	1305
		repeat_region	1248	1111	1306
		repeat_region	1249	1111	1307
		repeat_region	1250	1111	1308
		repeat_region	1251	1111	1309
		repeat_region	1252	1111	1310
		repeat_region	1253	1111	1311
		repeat_region	1254	1111	1312
		repeat_region	1255	1111	1313
		repeat_region	1256	1111	1314
		repeat_region	1257	1111	1315
		repeat_region	1258	1111	1316
		repeat_region	1259	1111	1317
		repeat_region	1260	1111	1318
		repeat_region	1261	1111	1319
		repeat_region	1262	1111	1320
		repeat_region	1263	1111	1321
		repeat_region	1264	1111	1322
		repeat_region	1265	1111	1323
		repeat_region	1266	1111	1324
		repeat_region	1267	1111	1325
		repeat_region	1268	1111	1326
		repeat_region	1269	1111	1327
		repeat_region	1270	1111	1328
		repeat_region	1271	1111	1329
		repeat_region	1272	1111	1330
		repeat_region	1273	1111	1331
		repeat_region	1274	1111	1332
		repeat_region	1275	1111	1333
		repeat_region	1276	1111	1334
		repeat_region	1277	1111	1335
		repeat_region	1278	1111	1336
		repeat_region	1279	1111	1337
		repeat_region	1280	1111	1338
		repeat_region	1281	1111	1339
		repeat_region	1282	1111	1340
		repeat_region	1283	1111	1341
		repeat_region	1284	1111	1342
		repeat_region	1285	1111	1343
		repeat_region	1286	1111	1344
		repeat_region	1287	1111	1345
		repeat_region	1288	1111	1346
		repeat_region	1289	1111	1347
		repeat_region	1290	1111	1348
		repeat_region	1291	1111	1349
		repeat_region	1292	1111	1350
		repeat_region	1293	111	

Qy 1771 CGAGCCCTTCAGGCCAGCAGCTGCCCCAGACTCGCTGGGCACTCAGTTCCCTCATCTGTA 1830
Db |||||||
Qy 45608 CGAGCCCTTCAGGCCAGCAGCTGCCCCAGACTCGCTGGGCACTCAGTTCCCTCATCTGTA 45667
Db |||||||
Qy 1831 AAGGTGAAGGGTGATGAGGATATGCTGACAGGAACAGTCTGTGGATGGACATGATCAG 1890
Db |||||||
Qy 45668 AAGGTGAAGGGTGATGAGGATATGCTGACAGGAACAGTCTGTGGATGGACATGATCAG 45727
Db |||||||
Qy 1891 TGCTNAAGGNAAGCAGCAGAGAGAGAGCGTTCGGCGCCCCGACGNCACCTNATCAGTGT 1950
Db |||||||
Qy 45728 TGCTAA--GGAAAGCAGCAGAGAGAGAGCGTTCGGCGCCCCGACGNCACCTNATCAGTGT-- 45783
Db |||||||
Qy 1951 NCCAGGTGCTNGGTTNCCCGAGNAGCAGAGCTNCAAGNATCANCACTGACACTNCAACC 2010
Db |||||||
Qy 45784 -CCAGCGTGTGGTTCCC-----CAGAGCAGAGCTCAGCATCACACTGACACTCACC 45834
Db |||||||
Qy 2011 TNGCCCTGCCCTNGCCANGAGGGTACTGCCGNAAGCGCACTTTGACACNTCTGATGNACC 2070
Db |||||||
Qy 45835 CTGCCCTGCCCTGG--CCAGAGGGTACTGCCG--ACGGCACTTTGACAC--TCTGATG--ACC 45889
Db |||||||
Qy 2071 TCAAAGCACTTTTCATGGCTNGCCCTCTNNGCAGGNCAGGNCAGGNCAGTGACANCT 2130
Db |||||||
Qy 45890 TCAAAGCACTTTTCATGGCTNGCCCTCTGCGAGGGCAGGGCAGGAGTGACACT----- 45942
Db |||||||
Qy 2131 GTAGNAGCATANGAANGCAGAGATGGGGTGNAAAGGGANCAACAGTCTTGAGCTGTCC 2190
Db |||||||
Qy 45943 ---GTAGGAGCATAGCAAGCAGAGATGGGGTGAA--GGGACACAGTCTTGAGCTGTCC 45997
Db |||||||
Qy 2191 ANCATGATGTGACTNCTCAAACTCTNNCCAGNATTTCTCTAAGAAATAGCANCCCC 2250
Db |||||||
Qy 45998 A-CATGATGTGACT-CCTCAAACTCTT---CCAGATTTCTCTAAGAAATAGCACCCCC- 46051
Db |||||||
Qy 2251 TTNCCCCATTTGCCCACTTAGCTTCTTCTCCAGGGGAGCTANCTCAGGACTCACGTAG 2310
Db |||||||
Qy 46052 -TTCCCCATTTGCCCACTTAGCTTCTTCTCCAGGGGAGCTA-CTCAGGACTCACGTAG 46109
Db |||||||
Qy 2311 CATTAATCAGCTGTGNAATGTCAGGGGGTGTCTGCTAGCCTCAACCTCTGGGGCAGG 2370
Db |||||||
Qy 46110 CATTAATCAGCTGTG-AATCGTCAGGGGGTGTCTGCTAGCCTCAACCTCTGGGGCAGG 46168
Db |||||||
Qy 2371 GGAGCCGAGACTCCGTGGGAGAGCTCATTCACATCTTGCCAGACAGCCTTTNGTC 2430
Db |||||||
Qy 46169 GGAGCCGAGACTCCGTGGGAGAGCTCATTCACATCTTGCCAGACAGCCTTT-GTC 46227
Db |||||||
Qy 2431 CAGCTGTCCATTTAGTTCAGACTGCTCCCGGGGAGAGAGCCCGGCCCCAGCACATAA 2490
Db |||||||
Qy 46228 CAGCTGTCCATTTAGTTCAGACTGCTCCCGGGGAGAGAGCCCGGCCCCAGCACATAA 46287
Db |||||||
Qy 2491 AGAACTCAGCCTTTGGTACTGACAGAGTCTGGGTTGTAGAGAACTCTTTTGAAGCAATAA 2550
Db |||||||
Qy 46288 AGAACTCAGCCTTTGGTACTGACAGAGTCTGGGTTGTAGAGAACTCTTTTGAAGCAATAA 46347
Db |||||||
Qy 2551 GTTTGGGGTGATGACAAATGTT 2572
Db |||||||
Qy 46348 GTTTGGGGTGATGACAAATGTT 46369
Db |||||||

Search completed: June 10, 2005, 14:22:48
Job time : 7500.59 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:57:41 ; Search time 30.0772 Seconds
(without alignments)
2390.553 Million cell updates/sec

Title: US-09-155-676B-2
Perfect score: 3093
Sequence: 1 XTGFGXMGSGSNXDKRQF.....FLMTFHVCEPINCFSHLKKK 604

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 557933 seqs, 119041758 residues

Total number of hits satisfying chosen parameters: 557933

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New:*

1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US11_NEW_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3058	98.9	604	5	US-09-155-676B-2
2	388	12.5	91	6	US-10-450-763-30582
3	302	9.8	520	7	US-11-097-143-11289
4	236	7.6	131	6	US-10-450-763-58372
5	193.5	6.3	1082	7	US-11-090-997-344
6	189	6.1	1094	7	US-11-090-997-1432
7	183.5	5.9	720	7	US-11-097-143-14316
8	181.5	5.9	269	8	US-60-655-875-154870
9	175.5	5.7	473	1	PCT-US03-10870-2491
10	171	5.5	1006	6	US-10-450-763-51537
11	160	5.2	366	8	US-60-669-175-21377
12	154	5.0	388	6	US-10-450-763-51533
13	151	4.9	600	8	US-60-680-002-931
14	151	4.9	601	8	US-60-671-501-404
15	143	4.6	183	6	US-10-450-763-51534
16	139.5	4.5	273	8	US-60-669-241-48091
17	136.5	4.4	2843	1	PCT-US03-10870-1970
18	136.5	4.4	2844	8	US-60-671-501-211
19	136	4.4	1139	6	US-10-940-774A-6556
20	136	4.4	1146	6	US-10-940-774A-7567
21	130.5	4.2	440	6	US-10-940-774A-9774
22	130	4.2	2829	6	US-10-525-621-1
23	123	4.0	972	8	US-60-643-717-15402
24	122	3.9	1945	7	US-11-097-143-21633
25	121.5	3.9	1035	7	US-11-097-143-2004

ALIGNMENTS

RESULT 1

US-09-155-676B-2
; Sequence 2, Application US/09155676B
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR
; FILE OF INVENTION: PREPARATION AND USE
; FILE REFERENCE: WALLACH-21
; CURRENT APPLICATION NUMBER: US/09/155.676B
; CURRENT FILING DATE: 1999-01-04
; PRIOR APPLICATION NUMBER: PCT/IL97/00117
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: IL 117800
; PRIOR FILING DATE: 1996-04-02
; PRIOR APPLICATION NUMBER: IL 119133
; PRIOR FILING DATE: 1996-08-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 604
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (6)..(6)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13)..(13)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15)..(15)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature

Sequence 1425, Ap
Sequence 10758, A
Sequence 10759, A
Sequence 2013, Ap
Sequence 45794, A
Sequence 54021, A
Sequence 34486, A
Sequence 1186, Ap
Sequence 29559, A
Sequence 1655, Ap
Sequence 12774, A
Sequence 23, Appl
Sequence 23, Appl
Sequence 10757, A
Sequence 42824, A
Sequence 47387, A
Sequence 1586, Ap
Sequence 943, App
Sequence 943, App
Sequence 943, App

Mon Jun 13 13:27:11 2005

LOCATION: (37)..(37)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (271)..(271)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (274)..(274)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (334)..(334)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (348)..(348)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (354)..(355)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (359)..(359)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (363)..(363)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (405)..(405)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (549)..(549)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (569)..(570)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid

US-09-155-676B-2
Query Match 98.9%; Score 3058; DB 5; Length 604;
Best Local Similarity 100.0%; Pred. No. 5.2e-221; Indels 0; Gaps 0;
Matches 603; Conservative 0; Mismatches 0;

QY 2 TGPXGXMGSXNKKRQFLERLLDAVKQCIRFXGRKEIASDSRVTCCLCAQFEAVL 61
DB 2 TGPXGXMGSXNKKRQFLERLLDAVKQCIRFXGRKEIASDSRVTCCLCAQFEAVL 61
QY 62 QHGLKRSRGLATAAAIKQAAGFASKTETEPVFWYVYKVELNKHQLORFVSLRHIAADVG 121
DB 62 QHGLKRSRGLATAAAIKQAAGFASKTETEPVFWYVYKVELNKHQLORFVSLRHIAADVG 121
QY 122 RGAWLRCALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERRSSMLPTMAAGLSILF 181
DB 122 RGAWLRCALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERRSSMLPTMAAGLSILF 181
QY 182 AINIDNDKLNQOSKFAPTVSDLLKESQNTVTSLLKESQGVSSLPREITASSAVSILIKP 241
DB 182 AINIDNDKLNQOSKFAPTVSDLLKESQNTVTSLLKESQGVSSLPREITASSAVSILIKP 241
QY 242 EQETDPCLSCPGMSVLPNNAKRSRRKRPKPTXSHLMRQMSRTLTCLKSHLGGQGAOR 301
DB 242 EQETDPCLSCPGMSVLPNNAKRSRRKRPKPTXSHLMRQMSRTLTCLKSHLGGQGAOR 301
QY 302 TTPTAPLISCPPLKAPSGLTPMSEQOLMENXEPVFERGVWVPEAXCEKHXRCXKX 361
DB 302 TTPTAPLISCPPLKAPSGLTPMSEQOLMENXEPVFERGVWVPEAXCEKHXRCXKX 361
QY 362 RXRVWKLIRKEAQGLGVAREATGREHPLPLDAQLGSAEGAAQXLRHPLPCQWRGLLOPS 421

DB 362 RXRVWKLIRKEAQGLGVAREATGREHPLPLDAQLGSAEGAAQXLRHPLPCQWRGLLOPS 421
QY 422 RCPPRKPGERDRTRGPRSPGWSVTCGSQLSRPKSSQPVTSASVPESMTISELROAT 481
DB 422 RCPPRKPGERDRTRGPRSPGWSVTCGSQLSRPKSSQPVTSASVPESMTISELROAT 481
QY 482 VAMNKRKDELEENRSLRNLLDGEHSAALROEVDTLKRKVAEQERQGMKVQALASYL 541
DB 482 VAMNKRKDELEENRSLRNLLDGEHSAALROEVDTLKRKVAEQERQGMKVQALASYL 541
QY 542 CYFVRFPXPHVRTMWRNGREKSNSSXXSHLSSWTQSFILKCLFMTFHVCEPINCFSHL 601
DB 542 CYFVRFPXPHVRTMWRNGREKSNSSXXSHLSSWTQSFILKCLFMTFHVCEPINCFSHL 601
QY 602 KKK 604
DB 602 KKK 604

RESULT 2

US-10-450-763-30582
; Sequence 30582, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 30582
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-30582

Query Match 12.5%; Score 388; DB 6; Length 91;
Best Local Similarity 96.3%; Pred. No. 7.4e-22;
Matches 79; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 MSGSXNKKRQFLERLLDAVKQCIRFXGRKEIASDSRVTCCLCAQFEAVLQHLKRS 68
DB 1 MSGSQNNDKROFLERLLDAVKQCIRFGGRKEIASDSRVTCCLCAQFEAVLQHLKRS 60
QY 69 RGLALTAATAAIKQAAGFASKTET 90
DB 61 RGLALTAATAAIKQAAGFASKTET 82

RESULT 3

US-11-097-143-11289
; Sequence 11289, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932

;; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
;; FILE REFERENCE: 38-21(53885)
;; CURRENT APPLICATION NUMBER: US/60/655.875
;; CURRENT FILING DATE: 2005-02-24
;; NUMBER OF SEQ ID NOS: 171306
;; SEQ ID NO 154870
;; LENGTH: 269
;; TYPE: PRT
;; ORGANISM: Heterodera glycines

;; NAME/KEY: misc feature
;; LOCATION: (241) (241)
;; OTHER INFORMATION: Xaa can be any naturally occurring amino acid

;; FEATURE:
;; OTHER INFORMATION: Coding regions on vcdna: vcdna-seqID 81293; Strands=-; Position=1
;; OTHER INFORMATION: -166,1016-1128,1249-1344,1396-1475,1529-1639,1721-1805,3272-3393
;; FEATURE:
;; OTHER INFORMATION: Homolog annotation: Hit_ID=NP_496906.2; Match level="QueryCovered
;; OTHER INFORMATION: =96%, HitCoverage=41%, E-value=2e-49, Identity=41%; Hit descrip
;; OTHER INFORMATION: =RUN domain containing protein (20146) [Caenorhabditis elegans]
;; OTHER INFORMATION: emb|CAA21666.2| Hypothetical protein Y51H1
US-60-655-875-154870

Query Match 5.9%; Score 181.5; DB 8; Length 269;
Best Local Similarity 26.9%; Pred. No. 8.9e-06;
Matches 60; Conservative 33; Mismatches 81; Indels 49; Gaps 8;

QY 27 DAVKQCOIR-----FXGRKEIASDSDSRVT-CLCAQFEAVLQHLKRSRGLATAAAI 78
DB 32 DVEKKCKERTESFLKTLUSGRK---SQISSELTQALCTVIEAVFHIGURDAFFL----- 81
QY 79 KQAGFASKTETEPFVYVYKVLNKHLELQRFYSLRHIAADVGRGR-----AWLRCA 130
DB 82 ---KGRSHRNRPSPNFPFVSKYQTSIKTQISLNLQIHTEIGRARAWIRVLNANRIV 138
QY 131 LNEHSLERYLHMLADRCIL-----STRYEDWSFVMDERSMPLPTMAAGLNSILFAINID 186
DB 139 LNEYSLDHYISMFKDTROMKQFYADDAFLRDS-----CVNSLGVH 184
QY 187 NKDLNGOSKEAPTVDLLKESTQNTVSL--LKESTQGVSSILR 227
DB 185 LKRINGLKITAPTNSLSLNTWTPSLSLAGLIKTAPSIGAFK 227

RESULT 9

PCT-US03-10870-2491
; Sequence 2491, Application PC/TUS0310870
; GENERAL INFORMATION:
; APPLICANT: Mitokor, Inc.
; APPLICANT: Buck Institute
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465PC
; CURRENT APPLICATION NUMBER: PCT/US03/10870
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2491
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-10870-2491

Query Match 5.7%; Score 175.5; DB 1; Length 473;
Best Local Similarity 28.5%; Pred. No. 5.2e-05;
Matches 75; Conservative 35; Mismatches 118; Indels 35; Gaps 12;

QY 5 GXGMSGXNXDKRQFLLERLLDAVKQCQIRFXGRK-----EIASDSDSRVTCLCAQF 57
DB 14 GCGGGGKXLSARNAVER-RNLITVC--RFSVKTLIDRSCTETIDDSPEFNNFAAIL 70
QY 58 EAVLQHLKR-----SRGLATAAAIK-QAAGFASKTETEPFVYVYKVLNKH-----HELQ 109
DB 71 EQILSHRLKEISQSCRWLAHLQIPLOGQVTFWG--YESPRSFMDYIRVACRKVSQNCICS 128
QY 110 FYSLRHIASDVGRGAWLRCAALNEHSLERYLHMLADRCILSTFYEDWSFVMDERSM 169
DB 129 IENWENVSSSRAGRAWIRVALMEKHLSEYIISTALRDFKTRRFYEDGAIVLGE--ANML 187
QY 170 PTMAAGLNSILFAINIDNKLNGOSKFPAPTVSDLLKESTQNTVSLK--ESTQGVSSILR 227
DB 188 AGMLGLNAIDFSCFKLKGEGLDGS---FPAVID-----YTPYLKYIQSSDSISSDEE 236
QY 228 EI-TASSAVSILIKPQETDPCL 249
DB 237 ELRTLGSSGSESTPENVGPPFL 259

RESULT 10

US-10-450-763-51537
; Sequence 51537, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 51537
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (957) (973)
; OTHER INFORMATION: Phorbol esters / diacylglycerol binding domain proteins
; OTHER INFORMATION: domain identified by eMATRIX, accession number BL004798, p-value
; OTHER INFORMATION: 7.429e-09, raw score of 12.57
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (476) (721)
; OTHER INFORMATION: PH domain identified by Pfam, accession name PH, E-value=2.5e
; OTHER INFORMATION: -09, Pfam score of 38.4
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) (1006)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-51537

Query Match 5.5%; Score 171; DB 6; Length 1006;
Best Local Similarity 22.7%; Pred. No. 0.00029;
Matches 103; Conservative 55; Mismatches 157; Indels 138; Gaps 20;

QY 83 GFASKTETEPFVYVYKVLNKHLELQRFYSLRHIAADVGRGAWLRCAALNEHSLERYLHM 142
DB 19 GFSSPAVQEANRGALCSCIIEHLEHLTFVNTDVGRCAWLRLALNDGLMECYLKL 78
QY 143 LLADRCILSTFYEDWSFVMDERSMPLPTMAAGLNSILFAINIDNKDLN-----CQSK 195
DB 79 LLQEQARLHEYOPTALLRDAEAGEFLLSFLOGLTSLSFYSKSAITLNEWTLPLALSG 138
QY 196 FAPTVDLLKESTQNTVSLKKESTQGVSSILF-----REITASS----- 233

```
Db 139 LCP-LSELDPSTSGAELQKESLDSISHSSGSEDIHVHSGHKKIRNQLTASSLSLDT 197
QY 234 AVSILIKPEQETDPCL-----SCPGMSVLMPNNAKESGRKRKXPTXSHL 277
Db 198 ASSQSCLNSOCLQENGSKSPDHCPEPMSCDSDLGTAEDSDSLQE----- 249
QY 278 MMRKMSRTLTGTLKRLHGQGAQRT---TPTAPLS-----ISCPLKAPSGLT PME 325
Db 250 VLLFESKAQVNSVPTN---GLSQETEIPTQASLSLHGLNTYTLHC---EAPAEPLPAQ 303
QY 326 SEQOLMENKFPVPERGVWPEAXCEKHXRCXKRRXVWKLIRKEAGPGVVAR--EA 383
Db 304 AASGTQD-----GVHVQEP-----RPOAPSLDLQOPVES 333
QY 384 TGREHLPLPDA-----OLGSAEGAAXLRHPLPCQWRGLLQPSRCPRKPGERDRTG 436
Db 334 TSGQQ---PSSTVSETARVGGQGLQKAQAH-----GAGLKLIVSS 373
QY 437 PRSP--GSWTSVQCGSQLSRPRKSSQPVTSAS 467
Db 374 PTSPRKNKMSIS---EDFEYRP--SREQPLESAS 401

RESULT 11
US-60-669-175-21377
; Sequence 21377, Application US/60669175
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; APPLICANT: Baum, James A
; APPLICANT: Gilbertson, Larry A
; APPLICANT: Kovalic, David K
; APPLICANT: LaRosa, Thomas J
; APPLICANT: Lu, Maolong
; APPLICANT: Munvikwa, Tichifa R. I.
; APPLICANT: Roberts, James K
; APPLICANT: Wu, Wei
; APPLICANT: Zhang, Bei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CONTROL OF INSECT INFESTATION IN PLANTS
; FILE REFERENCE: 38-21(53597)
; CURRENT APPLICATION NUMBER: US/60/669,175
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60560842
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 60565632
; PRIOR FILING DATE: 2004-04-27
; PRIOR APPLICATION NUMBER: 60579062
; PRIOR FILING DATE: 2004-06-11
; PRIOR APPLICATION NUMBER: 60603421
; PRIOR FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: 60617261
; PRIOR FILING DATE: 2004-10-11
; NUMBER OF SEQ ID NOS: 40774
; SEQ ID NO 21377
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Diabrotica virgifera
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID_1074; Strand=+; Position=317
; OTHER INFORMATION: Homolog annotation: Hit_ID=XP_395844.1; Match level="QueryCoverage=99%, HitCoverage=60%, E-value=4e-64, Identity=41%"; Hit description: "similar to CG6613-PA [Apis mellifera]"
; OTHER INFORMATION: Pfam annotation: Pfam ID=RUN; Match level="Score=108.7, E-value=1e-29, Copies=1"; Pfam description=RUN domain
US-60-669-175-21377

Query Match 5.2%; Score 160; DB 8; Length 366;
Best Local Similarity 23.0%; Pred. No. 0.00054;
Matches 62; Conservative 55; Mismatches 105; Indels 48; Gaps 9;
QY 18 RQFLERLLDAVKQCIQIRFXKKEATASDSDR-----VTCLCAQFEAVLQHLKRSRGL 71
```

```
Db 19 KKSITRQLSDCK--EIHVAGVEENKSPNFSEFSTTAATLCTAIEAIFLHGLRDT--- 73-
QY 72 ALTRAAAIKQAAGFASKTETEPVFWYVKEVLNKHLELQRFYSLRHIASDVGRGAWLRAL 131
Db 74 --LIHKFKKALADVE-QSEPSFWAPLLIIISHRQIIIEQITNLSQITTEVGQCRVIRLAL 130
QY 132 NEHSILERYLHMLLADRCRLSTFYEDWSFVMDERSSMLPTMAAGLNSILFPAINDKDLN 191
Db 131 NDCULSSYLMIRQDSSSLKSYKYKNAYVRDGLLDVAORLIEGVEAF----- 178
QY 192 GQSKFAPTVSDLLKESTQNTVTSLLKESTQGVSS-----LPREITASSAVSIL 238
Db 179 -KSFTLPNSSL--NTWPLPSLI---LSGLWSFTLSCPVAPCDVQAQSIITESMKYQ 232
QY 239 IKPQETDPCLSGPGMSVLMPNNAKRSRRK 268
Db 233 SNFNSSDTASIC---SYMNSQSOSGLRQ 259

RESULT 12
US-10-450-763-51533
; Sequence 51533, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 51533
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-51533

Query Match 5.0%; Score 154; DB 6; Length 388;
Best Local Similarity 25.7%; Pred. No. 0.0016;
Matches 53; Conservative 29; Mismatches 76; Indels 48; Gaps 6;
QY 95 WY----YVKEVLNKHLELQRFYSLRHIASDVGRGAWLRALNEHSILERYLHMLLADRCRLS 151
Db 53 WWRMDWTPLPSRHIIELEHLTFVNMVGRCAWLRLALNGLMECYLKLLEQEARLR 112
QY 152 TFYEDWSFVMDERSSMLPTMAAGLNSILFPAINDKDLN-----GQSKFAPTVSDLL 204
Db 113 EYQPTALLRDABEGEFLSLFQGLMSLSFELSYKSAILNEWTLTLPALSGLCP-LSELD 171
QY 205 KESTQNTVTSLLKESTQGV-----SSLFREITASSAVSIL 238
Db 172 PLSTSGAELQKESLDSISHSSGSEDIHVHSGHKKIRNOKPTASSLSLDTASSQLSCS 231
QY 239 IKPQETDPCLSGPGMSVLMPNNAKRS 264
Db 232 L-----NSDSCL-----LOENGSKS 246

RESULT 13
US-60-680-002-931
; Sequence 931, Application US/60680002
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: KIDNEY DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001601
; CURRENT APPLICATION NUMBER: US/60/680,002
; CURRENT FILING DATE: 2005-05-12
```



```
; NUMBER OF SEQ ID NOS: 2417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 931
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-680-002-931

Query Match
  4.9%; Score 151; DB 8; Length 600;
Best Local Similarity 21.3%; Pred. No. 0.0048;
Matches 111; Conservative 71; Mismatches 190; Indels 150; Gaps 20;

Qy 56 QFEAVLQHLKRSGLALTAATAAIKQAAGFASKTETEPVFWYVYKVLNKLHQLQRFYSLRH 115
Db 39 QFFVMEHCLK--HGLKVKKSFICQNKSPFGPLSELVEKLCPEASDIATS-----VRNLPE 91
Qy 116 IASDVGRGRAWLRCAALNEHSLERYLHMLLADRCLSTFYEDWSFVMDERSSMLPTMAAG 175
Db 92 LKTAVGRAWLWYALMOKKLADYLVKVIDNKHLLSEFYEPALMMEEE--GMVIVGLLVG 150
Qy 176 LNSILFAINIDNKDLNGQ-----SKFAPTVSDI--LKESTQNTVTSLLKESTQGVSSLFR 227
Db 151 LNVLDANLCLKGEDLDQSGVIDFSLYLKDVQDLGGKEH-ERITDVLQKNY-----202
Qy 228 EITASSAVSILIRPEQETDPCLSCPGMSVLPNNAKRSRRKRXKPTXSHLMRMKMSRTLG 287
Db 203 -----VEELNRHLSC---TVGDLQTKIDGLEK-----TNSKQLEELSAATDR 241
Qy 288 TCLKRHLGGQGAQTTPAPLSISCPPLKAPSGLTPMSEQQLMENKFPFVFERGVWVPE- 346
Db 242 IC-----SLOEQOQLREQNELIRERSEKSEVEI 269
Qy 347 -----AXCEKHXRCG-KRRXRVMKLRKEAQGPGLGVAREATGREHLPLPDAQLG-- 397
Db 270 TKQDTKVELEYTKQTQGLDEMYSVMKQLEKKEKVRLELEKEL-----ELQIGMK 320
Qy 398 -SAEGAAQXLRHLPQWGRGLLPSCRPPKPRGDRTRGPRSPGWSVTSVQCGSQLSRPR 456
Db 321 TEMEIAMKLL-----EKD-THEKQD-----TLVALRQOLEBEVK 352
Qy 457 KSSQOPVTSASVPESM-----TISELRQATVAMNRKDELEENRSLRNLLDGEHSA 510
Db 353 AINLQMFHKAQNAESSLQKNEAITSFEGKTNQVMSSMKQMEERLOHSEARARQGAERSH 412
Qy 511 ALRQEV-----DTLKRKV-AEQEERQGMK 533
Db 413 KLQELGGRIGALQQLSQLHEQCSSLEKELKSEKEQRAIQ 454

RESULT 14
US-60-671-501-404
; Sequence 404, Application US/60671501
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 51534
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-51534

Query Match
  4.6%; Score 143; DB 6; Length 183;
Best Local Similarity 28.6%; Pred. No. 0.0042;
Matches 53; Conservative 24; Mismatches 80; Indels 28; Gaps 5;

Qy 54 CAQFEAVLQHLKRSGLALTAATAAIKQAAGFASKTETETEPVFWYVYKVLNKLHQLQRFYSL 113
Db 20 CTPSTSKLPQGGK-----KKSA--HOKPLPQPVFWPLLKAVTPKHIISEHL 66
Qy 114 RHIASDVGRGRAWLRCAALNEHSLERYLHMLLADRCLSTFYEDWSFVMDERSSMLPTMA 173
Db 67 TFVMDVGRGRAWLRCAALNEHSLERYLHMLLADRCLSTFYEDWSFVMDERSSMLPTMA 115
Qy 174 AGLNSILFAINIDNKDLNGQSGKFAPTVSDI--LKESTQNTVTSLLKESTQGVSSLFRITASS 233

; NUMBER OF SEQ ID NOS: 2417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 931
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-680-002-931

Query Match
  4.9%; Score 151; DB 8; Length 601;
Best Local Similarity 21.3%; Pred. No. 0.0048;
Matches 111; Conservative 71; Mismatches 190; Indels 150; Gaps 20;

Qy 56 QFEAVLQHLKRSGLALTAATAAIKQAAGFASKTETEPVFWYVYKVLNKLHQLQRFYSLRH 115
```

Db 116 HSSGSEIEVHHGHKIRNQK--PTASSLSLDTAS--SSQLSCSLNCDSCLLQENGSKS 171
QY 234 AVSIL 238
Db 172 PDHVL 176

Search completed: June 7, 2005, 12:21:16
Job time : 31.0772 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:52:21 ; Search time 269.467 Seconds
(without alignments)
2618.037 Million cell updates/sec

Title: US-09-155-676B-2

Perfect score: 3093

Sequence: 1 XTPGXGXMGSXNXXDKRQF.....FLMTFHVCEBPINCFHSLKXK 604

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/1/paa/PCRUS COMB.pcp.*
- 2: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*
- 3: /cgn2_6/ptodata/1/paa/US07 COMB.pcp.*
- 4: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 5: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*
- 6: /cgn2_6/ptodata/1/paa/US10 COMB.pcp.*
- 7: /cgn2_6/ptodata/1/paa/US11 COMB.pcp.*
- 8: /cgn2_6/ptodata/1/paa/US12 COMB.pcp.*
- 9: /cgn2_6/ptodata/1/paa/US13 COMB.pcp.*
- 10: /cgn2_6/ptodata/1/paa/US14 COMB.pcp.*
- 11: /cgn2_6/ptodata/1/paa/US15 COMB.pcp.*
- 12: /cgn2_6/ptodata/1/paa/US16 COMB.pcp.*
- 13: /cgn2_6/ptodata/1/paa/US17 COMB.pcp.*
- 14: /cgn2_6/ptodata/1/paa/US18 COMB.pcp.*
- 15: /cgn2_6/ptodata/1/paa/US19 COMB.pcp.*
- 16: /cgn2_6/ptodata/1/paa/US20 COMB.pcp.*
- 17: /cgn2_6/ptodata/1/paa/US21 COMB.pcp.*
- 18: /cgn2_6/ptodata/1/paa/US22 COMB.pcp.*
- 19: /cgn2_6/ptodata/1/paa/US23 COMB.pcp.*
- 20: /cgn2_6/ptodata/1/paa/US24 COMB.pcp.*
- 21: /cgn2_6/ptodata/1/paa/US25 COMB.pcp.*
- 22: /cgn2_6/ptodata/1/paa/US26 COMB.pcp.*
- 23: /cgn2_6/ptodata/1/paa/US27 COMB.pcp.*
- 24: /cgn2_6/ptodata/1/paa/US28 COMB.pcp.*
- 25: /cgn2_6/ptodata/1/paa/US29 COMB.pcp.*
- 26: /cgn2_6/ptodata/1/paa/US30 COMB.pcp.*
- 27: /cgn2_6/ptodata/1/paa/US31 COMB.pcp.*
- 28: /cgn2_6/ptodata/1/paa/US32 COMB.pcp.*
- 29: /cgn2_6/ptodata/1/paa/US33 COMB.pcp.*
- 30: /cgn2_6/ptodata/1/paa/US34 COMB.pcp.*
- 31: /cgn2_6/ptodata/1/paa/US35 COMB.pcp.*
- 32: /cgn2_6/ptodata/1/paa/US36 COMB.pcp.*
- 33: /cgn2_6/ptodata/1/paa/US37 COMB.pcp.*
- 34: /cgn2_6/ptodata/1/paa/US38 COMB.pcp.*
- 35: /cgn2_6/ptodata/1/paa/US39 COMB.pcp.*
- 36: /cgn2_6/ptodata/1/paa/US40 COMB.pcp.*
- 37: /cgn2_6/ptodata/1/paa/US41 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	98.9	3058	604	15	US-09-155-676-2 Sequence 2, Appli
2	98.9	3058	604	15	US-09-155-676A-2 Sequence 2, Appli
3	1594.5	51.6	541	26	US-10-070-255-7 Sequence 7, Appli
4	1537.5	49.7	784	26	US-10-070-255-9 Sequence 9, Appli
5	1537.5	49.7	813	26	US-10-070-255-8 Sequence 8, Appli
6	1196	38.7	375	20	US-09-629-469A-14017 Sequence 14017, A
7	1196	38.7	375	35	US-10-917-503-14017 Sequence 14017, A
8	1194	38.6	250	27	US-10-170-205B-18638 Sequence 18638, A
9	721	23.3	320	22	US-09-758-472-5854 Sequence 5854, Ap
10	721	23.3	320	28	US-10-235-926-5854 Sequence 5854, Ap
11	693.5	22.4	548	37	US-60-212-356-236 Sequence 236, App
12	685.5	22.2	544	37	US-60-212-356-203 Sequence 203, App
13	663.5	21.5	692	37	US-60-229-525-294 Sequence 294, App
14	498.5	16.1	116	27	US-10-170-205B-22102 Sequence 22102, A
15	388	12.5	91	1	PCT-US01-08631-30582 Sequence 30582, A
16	374.5	12.1	424	27	US-10-170-205B-6178 Sequence 6178, Ap
17	348.5	11.3	329	37	US-60-229-525-456 Sequence 456, App
18	311	10.1	411	27	US-10-170-205B-6243 Sequence 6243, Ap
19	302	9.8	520	20	US-09-614-150-11289 Sequence 11289, A
20	302	9.8	520	20	US-09-614-150A-11289 Sequence 11289, A
21	302	9.8	520	37	US-60-167-217-11318 Sequence 11318, A
22	302	9.8	520	37	US-60-173-464-9160 Sequence 9160, Ap
23	302	9.8	520	37	US-60-191-637-11321 Sequence 11321, A
24	302	9.8	520	37	US-60-191-681-8867 Sequence 8867, Ap
25	242.5	7.8	75	27	US-10-170-205B-17822 Sequence 17822, A
26	236	7.6	131	1	PCT-US01-08631-58372 Sequence 58372, A
27	199.5	6.5	522	21	US-09-724-676-66567 Sequence 66567, A
28	199.5	6.5	522	21	US-09-724-676A-66567 Sequence 66567, A
29	183.5	5.9	720	20	US-09-614-150-14316 Sequence 14316, A
30	183.5	5.9	720	20	US-09-614-150A-14316 Sequence 14316, A
31	183.5	5.9	720	37	US-60-167-217-14393 Sequence 14393, A
32	183.5	5.9	720	37	US-60-191-637-14357 Sequence 14357, A
33	183	5.9	673	21	US-09-724-676-66566 Sequence 66566, A
34	183	5.9	673	21	US-09-724-676A-66566 Sequence 66566, A
35	183	5.9	788	21	US-09-724-676-66557 Sequence 66557, A
36	183	5.9	788	21	US-09-724-676-66558 Sequence 66558, A
37	183	5.9	788	21	US-09-724-676-66559 Sequence 66559, A
38	183	5.9	788	21	US-09-724-676-66560 Sequence 66560, A
39	183	5.9	788	21	US-09-724-676-66561 Sequence 66561, A
40	183	5.9	788	21	US-09-724-676-66562 Sequence 66562, A
41	183	5.9	788	21	US-09-724-676-66563 Sequence 66563, A
42	183	5.9	788	21	US-09-724-676-66564 Sequence 66564, A
43	183	5.9	788	21	US-09-724-676-66565 Sequence 66565, A
44	183	5.9	788	21	US-09-724-676A-66557 Sequence 66557, A
45	183	5.9	788	21	US-09-724-676A-66558 Sequence 66558, A

ALIGNMENTS

RESULT 1
US-09-155-676-2
; Sequence 2, Application US/09155676
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: MERT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMAR, P L L C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 04-JAN-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IL97/00117
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117800
; FILING DATE: 02-APR-1996
; APPLICATION NUMBER: IL 119133
; FILING DATE: 26-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: WALLACH-21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 604 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-155-676-2

Query Match 98.9%; Score 3058; DB 15; Length 604;
Best Local Similarity 100.0%; Pred. No. 4.2e-259;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TPGGXGMSGXNKKQFLLERLLDAVKQCOIRFXGRKEIASDSDSRVTCICAQFEAVL 61
Db 2 TPGGXGMSGXNKKQFLLERLLDAVKQCOIRFXGRKEIASDSDSRVTCICAQFEAVL 61

QY 62 QHGLKRSGLALTAATAIKQAAGFASKTETEPVFWYVYKVKELNKHQLQFYSLRHIASDVG 121
Db 62 QHGLKRSGLALTAATAIKQAAGFASKTETEPVFWYVYKVKELNKHQLQFYSLRHIASDVG 121

QY 122 RGRWLRCALNEHSLERYLHMLLADRCRLSTFYEDWMSFVMDDEERSMMLPTMAAGLNSILF 181
Db 122 RGRWLRCALNEHSLERYLHMLLADRCRLSTFYEDWMSFVMDDEERSMMLPTMAAGLNSILF 181

QY 182 AINIDNKLNGSKAPTVDLLKESTQNTVSLKESQGVSSLPREITASSAVSILIKP 241
Db 182 AINIDNKLNGSKAPTVDLLKESTQNTVSLKESQGVSSLPREITASSAVSILIKP 241

QY 242 EQETDPCILSCPGMSVLMFNKASGRKRKXPTXSHLMNKKMSRTLTGCKRHLGOGRAQR 301
Db 242 EQETDPCILSCPGMSVLMFNKASGRKRKXPTXSHLMNKKMSRTLTGCKRHLGOGRAQR 301

QY 302 TTPTAPLISICPPLKAPSGLTPTMSESQQLMENKXFPVFERGVWVPEAXCEKHXKXCGKXR 361
Db 302 TTPTAPLISICPPLKAPSGLTPTMSESQQLMENKXFPVFERGVWVPEAXCEKHXKXCGKXR 361

QY 362 RYVWVKLIRKEAQGLGVAREATGREHLPLDPAQLGSAAGAAQXLRHPLPCWRRGLLOPS 421
Db 362 RYVWVKLIRKEAQGLGVAREATGREHLPLDPAQLGSAAGAAQXLRHPLPCWRRGLLOPS 421

QY 422 RCPPRKPGERDRTRGPRSPGWSVQCGSQLSRPKSSQEPVTSASVPESMTIISLRQAT 481
Db 422 RCPPRKPGERDRTRGPRSPGWSVQCGSQLSRPKSSQEPVTSASVPESMTIISLRQAT 481

QY 482 VAMNMRKDELEENSLRNLLDGENMEHSAALRQEVDTLKRKVAQEEERQGMKVQALASYL 541
Db 482 VAMNMRKDELEENSLRNLLDGENMEHSAALRQEVDTLKRKVAQEEERQGMKVQALASYL 541

; 542 CYFVRRFXPHVRTMWRNGREKSNSSXXSHLSSWIOSFLKLCFLWTFHVCEPINCPSHL 601
; 542 CYFVRRFXPHVRTMWRNGREKSNSSXXSHLSSWIOSFLKLCFLWTFHVCEPINCPSHL 601
; 602 KKK 604
; 602 KKK 604

RESULT 2
US-09-155-676A-2
; Sequence 2, Application US/09155676A
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/155,676A
; FILING DATE: 04-JAN-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IL97/00117
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117800
; FILING DATE: 02-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 119133
; FILING DATE: 26-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: WALLACH-21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 604 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-155-676A-2

Query Match 98.9%; Score 3058; DB 15; Length 604;
Best Local Similarity 100.0%; Pred. No. 4.2e-259;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TPGGXGMSGXNKKQFLLERLLDAVKQCOIRFXGRKEIASDSDSRVTCICAQFEAVL 61
Db 2 TPGGXGMSGXNKKQFLLERLLDAVKQCOIRFXGRKEIASDSDSRVTCICAQFEAVL 61

QY 62 QHGLKRSGLALTAATAIKQAAGFASKTETEPVFWYVYKVKELNKHQLQFYSLRHIASDVG 121
Db 62 QHGLKRSGLALTAATAIKQAAGFASKTETEPVFWYVYKVKELNKHQLQFYSLRHIASDVG 121

QY 122 RGRWLRCALNEHSLERYLHMLLADRCRLSTFYEDWMSFVMDDEERSMMLPTMAAGLNSILF 181
Db 122 RGRWLRCALNEHSLERYLHMLLADRCRLSTFYEDWMSFVMDDEERSMMLPTMAAGLNSILF 181

QY 182 AINIDNKLNGSKAPTVDLLKESTQNTVSLKESQGVSSLPREITASSAVSILIKP 241
Db 182 AINIDNKLNGSKAPTVDLLKESTQNTVSLKESQGVSSLPREITASSAVSILIKP 241

QY 242 EQETDPCILSCPGMSVLMFNKASGRKRKXPTXSHLMNKKMSRTLTGCKRHLGOGRAQR 301
Db 242 EQETDPCILSCPGMSVLMFNKASGRKRKXPTXSHLMNKKMSRTLTGCKRHLGOGRAQR 301

QY 302 TTPTAPLISICPPLKAPSGLTPTMSESQQLMENKXFPVFERGVWVPEAXCEKHXKXCGKXR 361
Db 302 TTPTAPLISICPPLKAPSGLTPTMSESQQLMENKXFPVFERGVWVPEAXCEKHXKXCGKXR 361

QY 362 RYVWVKLIRKEAQGLGVAREATGREHLPLDPAQLGSAAGAAQXLRHPLPCWRRGLLOPS 421
Db 362 RYVWVKLIRKEAQGLGVAREATGREHLPLDPAQLGSAAGAAQXLRHPLPCWRRGLLOPS 421

QY 422 RCPPRKPGERDRTRGPRSPGWSVQCGSQLSRPKSSQEPVTSASVPESMTIISLRQAT 481
Db 422 RCPPRKPGERDRTRGPRSPGWSVQCGSQLSRPKSSQEPVTSASVPESMTIISLRQAT 481

QY 482 VAMNMRKDELEENSLRNLLDGENMEHSAALRQEVDTLKRKVAQEEERQGMKVQALASYL 541
Db 482 VAMNMRKDELEENSLRNLLDGENMEHSAALRQEVDTLKRKVAQEEERQGMKVQALASYL 541
```


Db 121 CALNEHSLERYLHMLADRCRLSTFYEDWDSFVMEERSSMLPTMAAGLSILFAINIDNK 180
QY 189 DLNGOSKAPFTVSDLLKKESTQNTVTSLLKESQGVSSLFREITASSAVSILIKPEQETDP 248
Db 181 DLNGOSKAPFTVSDLLKKESTQNTVTSLLKESQGVSSLFREITASSAVSILIKPEQETDP- 239
QY 249 LSCPGMSVLMNP-----AKRSGRKRKXPTXSHLMMRKMSRTLTGCLKRHLGQGRQAR- 301
Db 240 -----LPVVRNVSADAKCKKXKKTNNIISFDDDEEQNSGDVFKKTPGAGESSED 294
QY 302 -----TTPTAPLS-----ISCPLKAPSGLTPTME-----SEQOLMENX 334
Db 295 NSDRSSVNMISAFSPFGPNSGSSNSWKIDSLSLNGEFGYQKLDVKSIDDDVDENE 354
QY 335 FVPERGVWVPEACEKX-----XRCCKKRRXRVWKLIRKEAQGLGVARE- 382
Db 355 DDVY-----GNSGRKRGHSESPKPLEGNTC-----LSQMSWAPLKVHLND 398
QY 383 -----ATGREHLPLDPAQLGSAGAAQXLRHPLPCQWRGLLPQSRCPKPRGDRDRTRG 436
Db 399 SDILFPVSGVGSYPADAPLGSLNGTGPEDHVP-----DPLRYSVEA 443
QY 437 PRSPGWSVQCGSQLSRPKSSQPVTSASVPESMTISELRQATVAMNRRKDELEENR 496
Db 444 -SSPGH-----GSPLS-----SLLPSASVPESMTISELRQATVAMNRRKDELEENR 489
QY 497 SLRNLGEMEHSAAALRQEVDTLKRKVAEQERQGMKVOALA 538
Db 490 SLRNLGEMEHSAAALRQEVDTLKRKVAEQERQGMKVOALA 531

RESULT 5

US-10-070-255-8

; Sequence 8, Application US/10070255

; GENERAL INFORMATION:

; APPLICANT: WALLACH, David

; APPLICANT: MALININ, Nikolay

; APPLICANT: SINHA, Indranil

; APPLICANT: LEU, Stefan

; TITLE OF INVENTION: IREN PROTEIN, ITS PREPARATION AND USE

; FILE REFERENCE: WALLACH=28

; CURRENT APPLICATION NUMBER: US/10/070,255

; CURRENT FILING DATE: 2002-12-16

; PRIOR APPLICATION NUMBER: PCT/IL00/00517

; PRIOR FILING DATE: 2000-08-31

; PRIOR FILING DATE: 1999-09-02

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 8

; LENGTH: 813

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-070-255-8

Query Match 49.7%; Score 1537.5; DB 26; Length 813;
Best Local Similarity 61.2%; Pred. No. 4.2e-125;
Matches 356; Conservative 22; Mismatches 101; Indels 103; Gaps 13;

QY 9 MSGSXNXXKQRFLLERLLDAVKQCQIRFXGRKEIASDSRVTCLCAQFEAVLQHLKRS 68
Db 1 MSGSQNNKQRFLLERLLDAVKQCQIRFXGRKEIASDSRVTCLCAQFEAVLQHLKRS 60
QY 69 RGLALTAATIKQAAGFASKTETEPVFWYVKEVNLKHELQRFYSLRHIAADVGRGAWLR 128
Db 61 RGLALTAATIKQAAGFASKTETEPVFWYVKEVNLKHELQRFYSLRHIAADVGRGAWLR 120
QY 129 CALNEHSLERYLHMLADRCRLSTFYEDWDSFVMEERSSMLPTMAAGLSILFAINIDNK 188
Db 121 CALNEHSLERYLHMLADRCRLSTFYEDWDSFVMEERSSMLPTMAAGLSILFAINIDNK 180
QY 189 DLNGOSKAPFTVSDLLKKESTQNTVTSLLKESQGVSSLFREITASSAVSILIKPEQETDP 248

Db 181 DLNGOSKAPFTVSDLLKKESTQNTVTSLLKESQGVSSLFREITASSAVSILIKPEQETDP- 239
QY 249 LSCPGMSVLMNP-----AKRSGRKRKXPTXSHLMMRKMSRTLTGCLKRHLGQGRQAR- 301
Db 240 -----LPVVRNVSADAKCKKXKKTNNIISFDDDEEQNSGDVFKKTPGAGESSED 294
QY 302 -----TTPTAPLS-----ISCPLKAPSGLTPTME-----SEQOLMENX 334
Db 295 NSDRSSVNMISAFSPFGPNSGSSNSWKIDSLSLNGEFGYQKLDVKSIDDDVDENE 354
QY 335 FVPERGVWVPEACEKX-----XRCCKKRRXRVWKLIRKEAQGLGVARE- 382
Db 355 DDVY-----GNSGRKRGHSESPKPLEGNTC-----LSQMSWAPLKVHLND 398
QY 383 -----ATGREHLPLDPAQLGSAGAAQXLRHPLPCQWRGLLPQSRCPKPRGDRDRTRG 436
Db 399 SDILFPVSGVGSYPADAPLGSLNGTGPEDHVP-----DPLRYSVEA 443
QY 437 PRSPGWSVQCGSQLSRPKSSQPVTSASVPESMTISELRQATVAMNRRKDELEENR 496
Db 444 -SSPGH-----GSPLS-----SLLPSASVPESMTISELRQATVAMNRRKDELEENR 489
QY 497 SLRNLGEMEHSAAALRQEVDTLKRKVAEQERQGMKVOALA 538
Db 490 SLRNLGEMEHSAAALRQEVDTLKRKVAEQERQGMKVOALA 531

RESULT 6

US-09-629-469A-14017

; Sequence 14017, Application US/09629469A

; GENERAL INFORMATION:

; APPLICANT: OTA, TOSHIO

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: NISHIKAWA, TETSUO

; APPLICANT: HAYASHI, KOJI

; APPLICANT: SAITO, KAORU

; APPLICANT: YAMAMOTO, JUNICHI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: WAKAMATSU, AI

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: OTSUKI, TETSUJI

; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE

; FILE REFERENCE: 084335/0123

; CURRENT APPLICATION NUMBER: US/09/629,469A

; CURRENT FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: JP 1999-248036

; PRIOR FILING DATE: 1999-07-29

; PRIOR APPLICATION NUMBER: JP 1999-300253

; PRIOR FILING DATE: 1999-08-27

; PRIOR APPLICATION NUMBER: JP 2000-118776

; PRIOR FILING DATE: 2000-01-11

; PRIOR APPLICATION NUMBER: JP 2000-183767

; PRIOR FILING DATE: 2000-05-02

; PRIOR APPLICATION NUMBER: JP 2000-241899

; PRIOR FILING DATE: 2000-06-09

; PRIOR APPLICATION NUMBER: 60/159,590

; PRIOR FILING DATE: 1999-10-18

; PRIOR APPLICATION NUMBER: 60/183,322

; PRIOR FILING DATE: 2000-02-17

; NUMBER OF SEQ ID NOS: 19025

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 14017

; LENGTH: 375

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-629-469A-14017

Query Match 38.7%; Score 1196; DB 20; Length 375;
Best Local Similarity 82.2%; Pred. No. 1.4e-95;
Matches 244; Conservative 11; Mismatches 30; Indels 12; Gaps 2;
QY 9 MSGSXNXXKQRFLLERLLDAVKQCQIRFXGRKEIASDSRVTCLCAQFEAVLQHLKRS 68

Db 1 MSGSQNDKQFLERLLDAVKQCQIRFGGRKEIASDSRSKVTCLCAQFAVLQHGKRS 60
Qy 69 RGLALTAATAIKQAAGFASKTETEPVFWYVYKEVNLKHELQRFYSLRHIAADVGRGAWLR 128
Db 61 RGLALTAATAIKQAAGFASKTETEPVFWYVYKEVNLKHELQRFYSLRHIAADVGRGAWLR 120
Qy 129 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSSMLPTMAAGLSILFAINDNK 188
Db 121 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSSMLPTMAAGLSILFAINDNK 180
Qy 189 DLNGQSKFAPTVSDLLKESQNTVTSLLKESQGVSSLFREITASSAVSILIKPEQETDPC 248
Db 181 DLNGQSKFAPTVSDLLKESQNTVTSLLKESQGVSSLFREITASSAVSILIKPEQETDPC 239
Qy 249 LSCPGMSVLPN-----AKSGRRKRKXPTXSHLMRMKMSRTLGLTCLKRHLGQGRA 299
Db 240 -----LPVSRNVSADAKCKKXKVTNIIISFDDDEEQNSGDVFKTPTPGAGES 291

RESULT 7

US-10-917-503-14017
Sequence 14017, Application US/10917503

GENERAL INFORMATION:

APPLICANT: OTA, TOSHIO
APPLICANT: ISOGAI, TAKAO
APPLICANT: NISHIKAWA, TETSUO
APPLICANT: HAYASHI, KOJI
APPLICANT: SAITO, KAORU
APPLICANT: YAMAMOTO, JUNICHI
APPLICANT: ISHII, SHIZUKO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: WAKAMATSU, AI
APPLICANT: NAGAI, KEIICHI
APPLICANT: OTSUKI, TETSUJI

TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE

FILE REFERENCE: 084335/0123

CURRENT APPLICATION NUMBER: US/10/917,503

CURRENT FILING DATE: 2004-08-13

PRIOR APPLICATION NUMBER: US/09/629,469

PRIOR FILING DATE: 2000-07-28

PRIOR APPLICATION NUMBER: JP 1999-248036

PRIOR FILING DATE: 1999-07-29

PRIOR APPLICATION NUMBER: JP 1999-300253

PRIOR FILING DATE: 1999-08-27

PRIOR APPLICATION NUMBER: JP 2000-118776

PRIOR FILING DATE: 2000-01-11

PRIOR APPLICATION NUMBER: JP 2000-183767

PRIOR FILING DATE: 2000-05-02

PRIOR APPLICATION NUMBER: JP 2000-241899

PRIOR FILING DATE: 2000-06-09

PRIOR APPLICATION NUMBER: 60/159,590

PRIOR FILING DATE: 1999-10-18

PRIOR APPLICATION NUMBER: 60/183,322

PRIOR FILING DATE: 2000-02-17

NUMBER OF SEQ ID NOS: 19025

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 14017

LENGTH: 375

TYPE: PRT

ORGANISM: Homo sapiens

US-10-917-503-14017

Query Match 38.7%; Score 1196; DB 35; Length 375;

Best Local Similarity 82.2%; Pred. No. 1.4e-99;

Matches 244; Conservative 11; Mismatches 30; Indels 12; Gaps 2;

Qy 9 MSGSQNDKQFLERLLDAVKQCQIRFGGRKEIASDSRSKVTCLCAQFAVLQHGKRS 68
Db 1 MSGSQNDKQFLERLLDAVKQCQIRFGGRKEIASDSRSKVTCLCAQFAVLQHGKRS 60

Qy 69 RGLALTAATAIKQAAGFASKTETEPVFWYVYKEVNLKHELQRFYSLRHIAADVGRGAWLR 128
Db 1 MSGSQNDKQFLERLLDAVKQCQIRFGGRKEIASDSRSKVTCLCAQFAVLQHGKRS 60

Db 61 RGLALTAATAIKQAAGFASKTETEPVFWYVYKEVNLKHELQRFYSLRHIAADVGRGAWLR 120
Qy 129 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSSMLPTMAAGLSILFAINDNK 188
Db 121 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSSMLPTMAAGLSILFAINDNK 180
Qy 189 DLNGQSKFAPTVSDLLKESQNTVTSLLKESQGVSSLFREITASSAVSILIKPEQETDPC 248
Db 181 DLNGQSKFAPTVSDLLKESQNTVTSLLKESQGVSSLFREITASSAVSILIKPEQETDPC 239
Qy 249 LSCPGMSVLPN-----AKSGRRKRKXPTXSHLMRMKMSRTLGLTCLKRHLGQGRA 299
Db 240 -----LPVSRNVSADAKCKKXKVTNIIISFDDDEEQNSGDVFKTPTPGAGES 291

RESULT 8

US-10-170-205E-18638

Sequence 18638, Application US/10170205E

GENERAL INFORMATION:

APPLICANT: ADAMS, Mark

TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN CAPTURE AGENTS, AND USES THEREOF

FILE REFERENCE: CL001381

CURRENT APPLICATION NUMBER: US/10/170,205E

CURRENT FILING DATE: 2002-06-13

NUMBER OF SEQ ID NOS: 40312

SOFTWARE: Patent In version 3.2

SEQ ID NO 18638

LENGTH: 250

TYPE: PRT

ORGANISM: Homo sapiens

US-10-170-205E-18638

Query Match 38.6%; Score 1194; DB 27; Length 250;

Best Local Similarity 98.7%; Pred. No. 1.2e-95;

Matches 236; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 9 MSGSQNDKQFLERLLDAVKQCQIRFGGRKEIASDSRSKVTCLCAQFAVLQHGKRS 68
Db 1 MSGSQNDKQFLERLLDAVKQCQIRFGGRKEIASDSRSKVTCLCAQFAVLQHGKRS 60

Qy 69 RGLALTAATAIKQAAGFASKTETEPVFWYVYKEVNLKHELQRFYSLRHIAADVGRGAWLR 128
Db 61 RGLALTAATAIKQAAGFASKTETEPVFWYVYKEVNLKHELQRFYSLRHIAADVGRGAWLR 120

Qy 129 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSSMLPTMAAGLSILFAINDNK 188
Db 121 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSSMLPTMAAGLSILFAINDNK 180

Qy 189 DLNGQSKFAPTVSDLLKESQNTVTSLLKESQGVSSLFREITASSAVSILIKPEQETDPC 247
Db 181 DLNGQSKFAPTVSDLLKESQNTVTSLLKESQGVSSLFREITASSAVSILIKPEQETDPC 239

RESULT 9

US-09-758-472-5854

Sequence 5854, Application US/09758472

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PH001

CURRENT APPLICATION NUMBER: US/09/758,472

CURRENT FILING DATE: 2001-01-11

PRIOR APPLICATION NUMBER: 60/179,065

PRIOR FILING DATE: 2000-01-31

PRIOR APPLICATION NUMBER: 60/180,628

PRIOR FILING DATE: 2000-02-04

NUMBER OF SEQ ID NOS: 9632

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 5854

LENGTH: 320

TYPE: PRT

ORGANISM: Homo sapiens

```

; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (198)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (267)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (308)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (312)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-472-5854

Query Match      23.3%; Score 721; DB 22; Length 320;
Best Local Similarity 50.3%; Pred. No. 7.1e-54;
Matches 167; Conservative 21; Mismatches 64; Indels 80; Gaps 7;

QY 110 FYSLRHIASDVGRGAWLRCALNEHSRLERYLHMLLADRCRLSTFYEDWSFVMDERSMML 169
DB 19 FYSLRHIASDVGRGAWLRCALNEHSRLERYLHMLLADRCRLSTFYEDWSFVMDERSMML 78

QY 170 PTMAAGLNSILFAINIDKDLNGQSKFAPTVDLLKSTQNTVTSLLKSTQGVSSLFREI 229
DB 79 PTMAAGLNSILFAINIDKDLNGQSKFAPTVDLLKSTQNTVTSLLKSTQGVSSLFREI 138

QY 230 TASSAVSILIKPEQETDPCISCPGMSVLMNP-----AKSGRRKRXKPTXSHLMWRKMS 283
DB 139 TASSAVSILIKPEQETD-----LPVSVNRVNSADAKCKKRRKKKVTNIISFDDEDE 192

QY 284 RTLGTCLKHLGQGRAQRTTPTAPLSISCPPLKAPSLTPMSEQQLMENXFPVFERGVW 343
DB 193 QNSGDVFKKTPGAGESSE-----DNSDRSSVNIWSAFE-SPF 228

QY 344 VPEACEKHXRCXGKRRXRWKLRKEAQPLG-----VAREA 383
DB 229 GPNS-----NGSQSSNSWKIDSLSLNGEFGYQKLDVKSIDDEDDVYGN 278

QY 384 TGREHL-----PLPDAQLSAEG 401
DB 279 SGKRGHSESPRSHWKGTAPFPAQVGSXEG 310

RESULT 10
US-10-235-926-5854
; Sequence 5854, Application US/10235926
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PH001CIN
; CURRENT APPLICATION NUMBER: US/10/235,926
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/758,472
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 9632
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5854
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:

```

```

; NAME/KEY: misc feature
; LOCATION: (198)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (267)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (308)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-235-926-5854

Query Match      23.3%; Score 721; DB 28; Length 320;
Best Local Similarity 50.3%; Pred. No. 7.1e-54;
Matches 167; Conservative 21; Mismatches 64; Indels 80; Gaps 7;

QY 110 FYSLRHIASDVGRGAWLRCALNEHSRLERYLHMLLADRCRLSTFYEDWSFVMDERSMML 169
DB 19 FYSLRHIASDVGRGAWLRCALNEHSRLERYLHMLLADRCRLSTFYEDWSFVMDERSMML 78

QY 170 PTMAAGLNSILFAINIDKDLNGQSKFAPTVDLLKSTQNTVTSLLKSTQGVSSLFREI 229
DB 79 PTMAAGLNSILFAINIDKDLNGQSKFAPTVDLLKSTQNTVTSLLKSTQGVSSLFREI 138

QY 230 TASSAVSILIKPEQETDPCISCPGMSVLMNP-----AKSGRRKRXKPTXSHLMWRKMS 283
DB 139 TASSAVSILIKPEQETD-----LPVSVNRVNSADAKCKKRRKKKVTNIISFDDEDE 192

QY 284 RTLGTCLKHLGQGRAQRTTPTAPLSISCPPLKAPSLTPMSEQQLMENXFPVFERGVW 343
DB 193 QNSGDVFKKTPGAGESSE-----DNSDRSSVNIWSAFE-SPF 228

QY 344 VPEACEKHXRCXGKRRXRWKLRKEAQPLG-----VAREA 383
DB 229 GPNS-----NGSQSSNSWKIDSLSLNGEFGYQKLDVKSIDDEDDVYGN 278

QY 384 TGREHL-----PLPDAQLSAEG 401
DB 279 SGKRGHSESPRSHWKGTAPFPAQVGSXEG 310

RESULT 11
US-60-212-356-236
; Sequence 236, Application US/60212356
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE
; FILE REFERENCE: CL000677
; CURRENT APPLICATION NUMBER: US/60/212,356
; CURRENT FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 236
; LENGTH: 548
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(548)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-212-356-236

Query Match      22.4%; Score 693.5; DB 37; Length 548;
Best Local Similarity 56.7%; Pred. No. 4.3e-51;
Matches 160; Conservative 22; Mismatches 65; Indels 35; Gaps 5;

```


Qy	33	QIRFXGRKEIASDSDSRVTCICA-----QBPAVLQHGLKRSRGLALTAATAIKQAG	83
Db	135	QNOHTASLKIASILSSTEGGLCAVLDCWGDGEFATSAAASNVTSRRLAAVVLGSAIV	194
Qy	84	FASKTEPVPVYVVKVINKHELQRFYSLRHIASDVGRGAWLRCALNEHSLSERYLHML	143
Db	195	LWRESSTST-----SCKRFYSLRHIASDVGRGAWLRCALNEHSLSERYLHML	241
Qy	144	LADBCRLSTFYEDNSFVWDEERSMLPTMAAGLSILFAINIDNKDLNGQSKFAPTVDL	203
Db	242	LADBCRLSTFYEDNSFVWDEERSMLPTMAAGPNSILFAINIDNKDLNGQSKFAPTVDL	301
Qy	204	LKESTQNTVTLKKESTQGVSSLFREITASSAVSILIKPEQETDPCLSCPGMSVLMPN---	260
Db	302	LKESTQNTV--LLKESTQGVSSVFREITASSAISILIKPEQETDP-----LPVSRNVSA	354
Qy	261	---AKRGRRRKXKPTXSHLMWRKMSRTLGTCLKXHLGGRA	299
Db	355	DAKCKERKKKKQVNTNINSFDDEDEQNSGDMFKKTPGAGRS	396

```

RESULT 12
US-60-212-356-203
; Sequence 203, Application US/60212356
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO00677
; CURRENT APPLICATION NUMBER: US/60/212,356
; CURRENT FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203
; LENGTH: 544
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(544)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-212-356-203

```

```

RESULT 13
US-60-229-525-294
; Sequence 294, Application US/60229525
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF

```

```

; FILE REFERENCE: CL000772
; CURRENT APPLICATION NUMBER: US/60/229,525
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 819
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 294
; LENGTH: 692
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(692)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-229-525-294

```

```

RESULT 14
US-10-170-205E-22102
; Sequence 22102, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN
; ; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CU001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22102
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-22102

```

```

RESULT 15
PCT-US01-08631-30582
; Sequence 30582, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217

```

Mon Jun 13 13:27:10 2005

```
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 30582
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-08631-30582

Query Match      12.5%; Score 388; DB 1; Length 91;
Best Local Similarity 96.3%; Pred. No. 2e-25; 3; Indels 0; Gaps 0;
Matches 79; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      9 MSGSXNDKRFLLERLLDAVKQCQIRFXGRKEIASDSDSRVTCICAQFEAVLQHLKRS 68
Db      1 MSGSQNNDKRFLLERLLDAVKQCQIRFGGRKEIASDSDSRVTCICAQFEAVLQHLKRS 60

QY      69 RGLALTAATAIKOAGFASKTET 90
Db      61 RGLALTAATAIKOAGFASKTET 82

Search completed: June 7, 2005, 12:19:32
Job time : 271.467 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:20:34 ; Search time 5543.69 Seconds
(without alignments)
18065.048 Million cell updates/sec

Title: US-09-155-676B-3
Perfect score: 2631
Sequence: 1 ccccttcacagccagcc.....accctgacctcgagg99999 2631

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsl1:*
9: gb_gsl2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1770.4	67.3	3971	3	CR749592 Homo sapi
2	938.4	35.7	1640	3	CR612342 full-leng
3	872	33.1	2844	9	AY413577 Homo sapi
4	711.4	27.0	2765	9	AY413578 Pan trogl
5	706.2	26.8	911	5	BU195099 AGENCOURT
6	705	26.8	762	4	BI259434 602968032
7	698.6	26.6	4174	3	AK036200 Mus muscu
8	680.6	25.9	1140	4	BM458601 AGENCOURT
9	636.4	24.2	1133	1	AL556130
10	598.2	22.7	2829	9	AY413579 Mus muscu
11	568.6	21.6	702	7	CK979910 4111556 B
12	568.2	21.6	806	4	BI161002 602865208
13	564.2	21.4	580	5	BP267282 BP267282
c 14	553	21.0	1095	1	AL578455 AGENCOURT
15	514.2	19.5	955	5	BU541005 AGENCOURT
c 16	499.8	19.0	909	5	EX331085
17	456	17.3	1037	4	BG775354 602649808
18	446.8	17.0	773	7	CK365555 AGENCOURT
19	413.2	15.7	896	5	BU542238 AGENCOURT
20	408.8	15.5	991	4	BM550076 AGENCOURT
21	405.8	15.4	1072	5	BU517750 AGENCOURT
22	391	14.9	701	4	BG774889 602649808
23	389	14.8	768	4	BM048296 603625505
24	386	14.7	774	4	BG822015 602726122

25	385.2	14.6	819	5	BQ231411	AGENCOURT
26	379.8	14.4	692	4	BI651143	603297621
27	368	14.0	702	5	BI102204	EX102204
28	360.6	13.7	727	4	BI255165	602977452
c 29	348.4	13.2	545	7	N50047	YZ10e07.s1
30	343	13.0	615	5	BQ827984	LL6in2200
c 31	341.8	13.0	665	6	CB305346	UI-CF-EN1
32	340	12.9	744	6	CA321462	UI-M-FW0-
33	335.2	12.7	583	5	BP332147	BP332147
34	325.6	12.4	1172	4	BM543747	AGENCOURT
35	324.6	12.3	702	7	CN409540	170006001
36	322.2	12.2	579	2	BE407122	601301805
c 37	321.8	12.2	548	5	BU623799	UI-H-FG1-
38	319.8	12.2	702	4	BI489907	603031660
39	316.2	12.0	556	2	BE276256	601144348
c 40	315.2	12.0	531	6	CA748591	UI-H-FT1-
c 41	314.4	11.9	503	2	BE645094	7e65h06.x
c 42	312.8	11.9	454	1	AA807939	nu90f09.8
c 43	312.6	11.9	460	7	W65377	zd32d07.r1
c 44	311.8	11.9	445	1	AA621781	af06f05.s
c 45	309.2	11.8	450	4	BI685768	603309421

ALIGNMENTS

RESULT 1
CR749592
LOCUS CR749592 3971 bp mRNA linear HTC 19-AUG-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686J04131 (from clone DKFZp686J04131).
ACCESSION CR749592
VERSION CR749592.1 GI:51476844
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3971)
AUTHORS Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
CONSRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp686J04131) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686J04131
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES
source
Location/Qualifiers
1..3971
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686J04131"
/tissue_type="uterus"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="MAP3K14 protein, N-terminus truncated, differentially spliced"
1..3971
/gene="DKFZp686J04131"
<1..1302
/gene="DKFZp686J04131"
/codon_start=1
/product="hypothetical protein"

Db	3293	TCCGGCGCCCGAGCCCCCACTATCAGTGT---CCAGCGTGTGTTGTTCCC-----CAGA	3344		
Qy	1981	GCTNCAGNCATCANCACACTGACACACTNCACACCTTNGCCCTGCCCTTNGGCCAGANGAGGTA	2040		
Db	3342	GCACAGCTCAGCATCACACTGACACTCACCCTGCCCCTGCCCTCG--CCAGAGGGTA	3399		
Qy	2041	CCGNAGGCACCTTTGCACNTCTGATGNACCTCAAAGCACTTTTCATGGCTNGCCCTCTNN	2100		
Db	3400	CCG-ACGGCACTTTGCACT--CTGATGACCTCAAAGCACTTTTCATGGCTGCTCTTGCCA	3456		
Qy	2101	GCAGGNCAGGNCAGGNCAGTGACACTCTGTAGNAGCATANGCAANGCCAGAGATGG	2160		
Db	3457	GGCAGGCGAGGCGAGTGACACT-----GTAGAGCATAGCAAGCCAGAGATGG	3506		
Qy	2161	GGTGNAAAGGANCACAGTCTTTGAGCTGTCCANCATGCTATGCTGACTNCTCAAACTCTTN	2220		
Db	3507	GGTGAA--GGGACACAGTCTTTGAGCTGTCCA-CATGCATGTGACT-CCTCAACCTCTT-	3561		
Qy	2221	NCCAGNATTTCTTAAGAATAGNANCCCCCTTNGCCCATTTGCCGCCAGCTTAGCTCTTCT	2280		
Db	3562	--CCAGATTTCTTAAGAATAGCACCCCC--TTCCCCATTGCCCCAGCTTAGCTCTTCT	3617		
Qy	2281	CCCAGGGGAGCTANCTCAGGACTCAGCTGAGCATTAATCAGCTGTGNAATCGTCAGGGGG	2340		
Db	3618	CCCAGGGGAGCTA-CTCAGGACTCAGTGAAGCATTAATCAGCTGTG-AACTCGTCAGGGGG	3675		
Qy	2341	TGTTCTGTAGCTCAACCTCTCTGGGGCAGGGGAGCGGAGACTCCGTGGGAGAGGCTCAT	2400		
Db	3676	TGTTCTGTAGCTCAACCTCTCTGGGGCAGGGGAGCGGAGACTCCGTGGGAGAGGCTCAT	3735		
Qy	2401	TCCCATCTTTGCCAAGACAGCCTTTTNGTTCAGCTGTGCCATTTGAGTCAAGCTGTCTCC	2460		
Db	3736	TCCCATCTTTGCCAAGACAGCCTTT-TGCCAGCTGTCCATTTGAGTCAAGCTGTCTCC	3794		
Qy	2461	GGGGAGAGAGCCCGGGCCCCCAGACATAAAGAACTGCAGCTTGGTACTGCAGAGCTG	2520		
Db	3795	GGGGAGAGAGCCCGGGCCCCCAGACATAAAGAACTGCAGCTTGGTACTGCAGAGCTG	3854		
Qy	2521	GGTTGTAGAGAACTCTTTGTAAAGCAATAAGCTTTGGGGTGTATGACAAATGTTAAAAA	2579		
Db	3855	GGTTGTAGAGAACTCTTTGTAAAGCAATAAGCTTTGGGGTGTATGACAAATGTTAAAAA	3913		
RESULT 2	CR612342	1640 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	full-length cDNA clone CS0DK011Y14 of Hela cells Cot 25-normalized				
DEFINITION	of Homo sapiens (human).				
ACCESSION	CR612342				
VERSION	CR612342.1	GI:50493149			
KEYWORDS	HTC; CNSLT cDNA.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 1640)				
JOURNAL	Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.				
REMARK	Full-length cDNA libraries and normalized				
	Unpublished				
	Contact : Feng Liang Email : fliang@lifetech.com URL :				
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600				
	Faraday Avenue				
	2 (bases 1 to 1640)				
REFERENCE	Genoscope.				
AUTHORS	Direct Submission				
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequences :				
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
	- Web : www.genoscope.cns.fr)				
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime				
	end enriched, double-strand cDNA was digested with Not I and cloned				
	into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library				
	was normalized. Library was constructed by Life Technologies, a				
	division of Invitrogen.				

```
QY 1724 GTGCCCTTTGCTGCTGCTACCACTCTTCCCAGAGCAGAGGCCCGAGCCCTTCCAG 1783
Db 854 GTGCCCTTTGCTGCTGCTGCTACCACTCTTCCCAGAGCAGAGGCCCGAGCCCTTCCAG 913
QY 1784 GCCCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTCTGTAAGGTGAAGGGTG 1843
Db 914 GCCCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTCTGTAAGGTGAAGGGTG 973
QY 1844 ATGCAGGATATGCTGTCAGAGAACAGTCTGTGGATGGACATGATCAGTGTCTTAAAGNAAA 1903
Db 974 ATGCAGGATATGCTGTCAGAGAACAGTCTGTGGATGGACATGATCAGTGTCTTAAAGNAAA 1031
QY 1904 GCAGCAGAGAGAGAGCTGCTCGGGGCCCGAGNCCCAGCTNATCAGTGTNCCAGCGTCTG 1963
Db 1032 GCAGCAGAGAGAGAGCTGCTCGGGGCCCGAGNCCCAGCTNATCAGTGTNCCAGCGTCTG 1088
QY 1964 GTTNCCTCCAGNAGCAGACTNCCAGNCCATCANCACTGACACTNCCAGTGTAGNAGCATAN 2023
Db 1089 TTCCCC-----CAGACACAGCTCAGCATCACACTGACACTCACCCTGCCCCCT 1140
QY 2024 NGGCCANGAGGTACTGCGCGNAGGCACTTTGCACTNCTGTATGNACCTCAAGACATTTTC 2083
Db 1141 GG--CCAGAGGCTACTGCGG--ACGGCACTTTGCAC--TCTGATG--ACCTCAAGACATTTTC 1195
QY 2084 ATGGCTNGCCCTCTNNCGAGGNCAGGNCAGGNCAGTGCACANCTGTAGNAGCATAN 2143
Db 1196 ATGGCTTGCCCTCTGCGCAGGCGAGGCGAGGCGAGTGCACCT-----GTAGAGCAT 1245
QY 2144 GCAAGCCAGGAGATGGGGTGNAAAGGANCACAGTCTTGAGCTGTCCANCACTGATGA 2203
Db 1246 AGCAAGCCAGGAGATGGGGTGA--GGGACACAGTCTTGAGCTGTCCA--CATGCACTGA 1302
QY 2204 CTNCCCTCAAACCTCTTNNCCAGNATTTCTTAAGATAGCAGNCCCTTNCCTTCCCTG 2263
Db 1303 CT--CCTCAAAACCTCTT---CCAGATTCTCTAAGATAGCAGCCCC--TTCCCCCATG 1356
QY 2264 CAGACTTAGCTCTTCTCCAGGGAGTANCTCAGGACTCAGTGTAGCATTAATCAGCT 2323
Db 1357 CCAGCTTAGCTCTTCTCCAGGGAGCTA--CTCAGACTCAGTGTAGCATTAATCAGCT 1415
QY 2324 GTGNAATCGTCAGGGGTGTCTGTAGCTCTCACTCTCGGGCAGGGACGCGAGACT 2383
Db 1416 GTG--AATCGTCAGGGGTGTCTGTAGCTCTCACTCTCGGGCAGGGACGCGAGACT 1474
QY 2384 CCGTGGAGAGCTCAATCCACATCTTGCCAGACAGCTTTNGTCCAGCTGTCCACAT 2443
Db 1475 CCGTGGAGAGCTCAATCCACATCTTGCCAGACAGCTTT--GTCCAGCTGTCCACAT 1533
QY 2444 TGAGTCAGACTGCTCCGGGGAGAGAGCCCGGCCCGCCAGCACATAAAGACTGCGAGCT 2503
Db 1534 TGAGTCAGACTGCTCCGGGGAGAGAGCCCGGCCCGCCAGCACATAAAGACTGCGAGCT 1593
QY 2504 TGCTACTGTCAGAGTCTGGTGTGTAGAGAACTTTTGTAAAGCAATAAA 2550
Db 1594 TGCTACTGTCAGAGTCTGGTGTGTAGAGAACTTTTGTAAAGCAATAAA 1640

RESULT 3
AY413577 2844 bp DNA linear GSS 17-DEC-2003
LOCUS Homo sapiens MAP3K14 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY413577
VERSION AY413577.1 GI:39769539
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 2844)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2. (bases 1 to 2844)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..2844
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1_>2844
/gene="MAP3K14"
/locus_tag="HCM4937"
gene
ORIGIN
Query Match 33.1%; Score 872; DB 9; Length 2844;
Best Local Similarity 100.0%; Pred. No. 3.8e-216;
Matches 872; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 106 TGGGAGGTCTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACGCCAA 165
Db 1973 TGGGAGGTCTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACGCCAA 2032
QY 166 ATCAGCCCAATTACACACAGACCCCTCCATGCCAGCCGAGAGAGCTTTGCGCAAGGGCCC 225
Db 2033 ATCAAGCCCAATTACACACAGACCCCTCCATGCCAGCCGAGAGAGCTTTGCGCAAGGGCCC 2092
QY 226 CAGGCCCCGGCCAGCTGAGGAGACAAACAGGACAGCCCTTAAGCTCAGAGCTCTCTCC 285
Db 2093 CAGGCCCCGGCCAGCTGAGGAGACAAACAGGACAGCCCTTAAGCTCAGAGCTCTCTCC 2152
QY 286 CACAGAGCCCCCAGAGCCAAACAAAGTCTCTCCCTTCACTTTGAGCAGAGGAGTCTG 345
Db 2153 CACAGAGCCCCCAGAGCCAAACAAAGTCTCTCCCTTCACTTTGAGCAGAGGAGTCTG 2212
QY 346 GGATCTGGGACCCCTTACTCTCTCTGAGCCAGCCCTCCAGAAACCCAGCT 405
Db 2213 GGATCTGGGACCCCTTACTCTCTCTGAGCCAGCCCTCCAGAAACCCAGCT 2272
QY 406 CACAGAGCGGAAACAAACCGTCCCGAGCAGGAACTGCAGAGCTGGAAATAGAAATTAT 465
Db 2273 CACAGAGCGGAAACAAACCGTCCCGAGCAGGAACTGCAGAGCTGGAAATAGAAATTAT 2332
QY 466 TCTCAACAGCTGTCTCCAGCCATTTCTCTGAGGAGCAGAGCAAAATTTCTCTGGTGC 525
Db 2333 TCTCAACAGCTGTCTCCAGCCATTTCTCTGAGGAGCAGAGCAAAATTTCTCTGGTGC 2392
QY 526 TCAGCATCGACAGCTCTCCCTGTCGGATGACAGTGAGAAAGAAACCCATCAAGAGCTCTC 585
Db 2393 TCAGCATCGACAGCTCTCCCTGTCGGATGACAGTGAGAAAGAAACCCATCAAGAGCTCTC 2452
QY 586 AAAGCTCGCGGACACCCCTGAGCTCAGGCGTACACTCTTGGAGAGCAGCGCGAGGCTC 645
Db 2453 AAAGCTCGCGGACACCCCTGAGCTCAGGCGTACACTCTTGGAGAGCAGCGCGAGGCTC 2512
QY 646 GAAGCTCCAGCTGGAACATGGTGTGCGCCCGGGGGGGCCCGACCCAGCACTAT 705
Db 2513 GAAGCTCCAGCTGGAACATGGTGTGCGCCCGGGGGGGCCCGACCCAGCACTAT 2572
QY 706 TCAATGGTGTGAAGTCCCAATACAGTCTCTTAATGGTGAACACCTTGCACTCCGGAGT 765
Db 2573 TCAATGGTGTGAAGTCCCAATACAGTCTCTTAATGGTGAACACCTTGCACTCCGGAGT 2632
```

```
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2. (bases 1 to 2844)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..2844
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1_>2844
/gene="MAP3K14"
/locus_tag="HCM4937"
gene
ORIGIN
```

```

QY 766 TCACCGGGTCAAAGTGGGAGACATCGCCACTGGGATCAGAGCCAGATCCAGCTGCAG 825
DB 2633 TCACCGGGTCAAAGTGGGAGACATCGCCACTGGGATCAGAGCCAGATCCAGCTGCAG 2692
QY 826 CCTTCAGCTTGGTCAACAAAGAGCGGAGCGCTGTTTCGCTAGACATGAGGTGCCAGACT 885
DB 2693 CCTTCAGCTTGGTCAACAAAGAGCGGAGCGCTGTTTCGCTAGACATGAGGTGCCAGACT 2752
QY 886 CGGGCATCGACCTGACGTGCACACCTGGCCCTCGATGGCAGCTTCGCTGGAGCTGGAGGG 945
DB 2753 CGGGCATCGACCTGACGTGCACACCTGGCCCTCGATGGCAGCTTCGCTGGAGCTGGAGGG 2812
QY 946 TCAAGCATGGCCAGCTGGAGAACAGGCCCTTAA 977
DB 2813 TCAAGCATGGCCAGCTGGAGAACAGGCCCTTAA 2844

RESULT 4
AY413578 2765 bp DNA linear GSS 17-DEC-2003
LOCUS Pan troglodytes MAP3K14 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY413578
VERSION AY413578.1 GI:39769540
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 2765)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2765)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..2765
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>2765
/gene="MAP3K14"
/locus_tag="HCW4937"

gene
Query Match 27.0%; Score 711.4; DB 9; Length 2765;
Best Local Similarity 82.0%; Pred. No. 3.7e-174;
Matches 715; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 106 TGGGAGGTCAGAGCCCTTGGAGGGGAGAAATATAGAGACCAAGACATCCAGCCCAA 165
DB 1894 TGGGAGGTCAGAGCCCTTGGAGGGGAGAAATATAGAGACCAAGACATCCAGCCCAA 1953
QY 166 ATCAAGCCAAATTACCAACAGACCTTCATGCCAGCCGAGAGAGCTTTCGCCAAGGGGCC 225
DB 1954 ATCAAGCCAAATTATCACCAGACCTTCATGCCAGCCGAGAGAGCTTTCGCCAAGGGGCC 2013
QY 226 CAGGCCCGCCAGCTGAGGAGACAAACAGCGAGGCCCTTAAGCTCAGGCTCCTCTCC 285

```

```

Db 2014 CAGGCCCGCCAGCTGAGGAGACAACAGCGAGAGCCCTTAAGCTCAGGCTCCTCTCC 2073
QY 286 CACCAGAGCCCCAGAGCCAAACAAGTCTCTCCTCTGACTTTTGAGCAAGAGAGTCTG 345
Db 2074 CACCANAGCCCCAGAGCCAAACAAGTNTCTCCTCTGACTTTTGAGCAAGAGAGTCTG 2133
QY 346 GGATGTGGAAACCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 405
Db 2134 GGATGTGGAAACCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2193
QY 406 CACCAGAGCGAAAGAACCGTCCGGAGCAGGAACCTGACAGCTGGAATAGAAATAT 465
Db 2194 CACCAGAGCGAAAGAACCGTCCGGAGCAGGAACCTGACAGCTGGAATAGAAATAT 2253
QY 466 TCCTCAACAGCTGTCTCCAGCCATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 525
Db 2254 TCCTCAACAGCTGTCTCCAGCCATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2313
QY 526 TCAGCATCGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 585
Db 2314 TCAGCATCGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2373
QY 586 AAAGCTCGCGGACACCTCTGAGCTCAGCGTACACTCTCTCTCTCTCTCTCTCTCTCTCT 645
Db 2374 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2433
QY 646 GAAGCTCCAGCTGGAACATGTGTCTGTGGCCCGGGCGGCCACCGACACCCCAAGCTATT 705
Db 2434 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2493
QY 706 TCAATGGTGTGAAAGTCCAAATACAGTCTCTTAATGGTGAAACACTCTGCACATCCGGAGT 765
Db 2494 NNNNNNGTGTGAAAGTCCAAATACAGTCTCTTAATGGTGAAACACTCTGCACATCCGGAGT 2553
QY 766 TCCACCGGTCAAAGTGGGAGACATCCGCTGGCATCAGCAGCAGATCCAGCTGCAG 825
Db 2554 TCCACCGGTCAAAGTGGGAGACATCCGCTGGCATCAGCAGCAGATCCAGCTGCAG 2613
QY 826 CTTTCAGCTTGGTCAACAAAGACGGGAGCGCTGTTCTGCTACGATGAGGTGCCAGACT 885
Db 2614 CTTTCAGCTTGGTCAACAAAGACGGGAGCGCTGTTCTGCTACGATGAGGTGCCAGACT 2673
QY 886 CGGCGATCGACCTGACATGTCACATCTGGCCCTCTGATGGCAGCTTCGCTTGGAGCTGGAGGG 945
Db 2674 CGGCGATCGACCTGACATGTCACATCTGGCCCTCTGATGGCAGCTTCGCTTGGAGCTGGAGGG 2733
QY 946 TCAGCATGGCCAGCTGGAGAACAGGCCCTTAA 977
Db 2734 TCAGCATGGCCAGCTGGAGAACAGGCCCTTAA 2765

```

```

RESULT 5
BUI95099 911 bp mRNA linear EST 04-SEP-2002
LOCUS AGENCOURT 7964504 NIH_MGC_72 Homo sapiens cdna clone IMAGE:6169272
DEFINITION 5', mRNA sequence.
ACCESSION BUI95099
VERSION BUI95099.1 GI:22709083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 911)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

```


Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1534 row: c column: 01
High quality sequence stop: 556.
Location/Qualifiers

FEATURES

source

```
1. .911
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6169272"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."
```

ORIGIN

```
Query Match          26.8%; Score 706.2; DB 5; Length 911;
Best Local Similarity 91.3%; Pred. No. 5.9e-173;
Matches 837; Conservative 5; Mismatches 58; Indels 17; Gaps 9;

QY 420 GCAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAATTATTCTCAACAGCCTG 479
DB 1 GCACCGTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAATTATTCTCAACAGCCTG 60
QY 480 TCCAGCCATTCTCTGAGGAGCAGGAGCAAAATCTCTCTGCTCTAGCATCGACAGC 539
DB 61 TCCAGCCATTCTCTGAGGAGCAGGAGCAAAATCTCTCTGCTCTAGCATCGACAGC 120
QY 540 CTCTCCCTCGGATGACAGTGGAGAACCCATCAAGGCTCTCAAGCTCGGGGAC 599
DB 121 CTCTCCCTCGGATGACAGTGGAGAACCCATCAAGGCTCTCAAGCTCGGGGAC 180
QY 600 ACCCTGAGCTCAGGCTTACACTCTCTGAGCAGCCAGCCGAGCTCGAAGCTCCAGCTGG 659
DB 181 ACCCTGAGCTCAGGCTTACACTCTCTGAGCAGCCAGCCGAGCTCGAAGCTCCAGCTGG 240
QY 660 AACATGGTGTGCGCCCGGGGGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAA 719
DB 241 AACATGGTGTGCGCCCGGGGGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAA 300
QY 720 GTCCAAATACAGTCTCTTAATGTGAACACCTGCACATCGGAGGTTCACCGGGTCAAA 779
DB 301 GTCCAAATACAGTCTCTTAATGTGAACACCTGCACATCGGAGGTTCACCGGGTCAAA 360
QY 780 GTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCGAGCTGCAGCTTCAGCTTGGTC 839
DB 361 GTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCGAGCTGCAGCTTCAGCTTGGTC 420
QY 840 ACCAAAGAGCGGAGCCTGTTTGCTACGACATGGAGGTGCGACACTCGGGGATCGACCTG 899
DB 421 ACCAAAGAGCGGAGCCTGTTTGCTACGACATGGAGGTGCGACACTCGGGGATCGACCTG 480
QY 900 CAGTGACACATGGCCCTCATGTGCAGCTTCGCTGGAGCTGGAGGTTCAGGATGCCACG 959
DB 481 CAGTGACACATGGCCCTCATGTGCAGCTTCGCTGGAGCTGGAGGTTCAGGATGCCACG 540
QY 960 CTGGAGAACAGGCGCTTAACCTCTCCACCGCCGGCTCCACACTCGCGGAAAGCAGCC 1019
DB 541 CTGGAGAACAGGCGCTTAACCTCTCCACCGCCGGCTCCACACTCGCGG -AAGCAGCC 599
QY 1020 TTCTCTGCTGGTGCAGATGCTGCTCTGAAACACAGGCTCAGCGCTTCGAGGGGATTT 1079
DB 600 TTCTCTGCTGGTGCAGATGCTGCTCTGAAACACAGGCTCAGCGCTTCGAGGGGATTT 658
QY 1080 GNCACGCCCGCGCTCARGAGNTGGAAACGAGGCGCTCGNACGAGNAGAGTNGGG 1139
DB 659 GGCAG -CCCCCGGCTCAGCAGTGGGACAGGCGCTCGAGCAGC-----AAGTGGGG 712
QY 1140 GCAAGCNAGATGCTCTCCAGGATTTTCAACNCTGAGCCCTGCCCCCANCCTGCTGAAD 1199
```

```
DB 713 GCAACGAGGAATGCTCCCGAGATTTACACCCCTGAGCCT -GCCACACCTGGCTGAG 770
QY 1200 AAAACAYTNCGCCACAGTGAAGACAGAGGAGATGNCAGGAGTTNNACCTYGGGA 1259
DB 771 AAAACACTCCGCCACAGTGAAGACAGAGGAGTGAAGAGTTA--CCTCGGG 826
QY 1260 AACAAACAGGAGTCTTTTCTGCCCCTGCTCCAGTNCAGATT-GGCTTGNACCCGCTT 1318
DB 827 AAAACACAGGAGTCTT--CTCTGCCCTGCTTCAAGTCGAGTTGGGCTGACCCCGCTT 884
QY 1319 GGANTCAGTGACCAATT 1335
DB 885 TGGATCCGTGACCAATT 901

RESULT 6
BI259434
LOCUS
DEFINITION
602968032F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5107807 5',
mRNA sequence.
ACCESSION
BI259434
VERSION
BI259434.1 GI:14816769
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 762)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1260 row: o column: 08
High quality sequence stop: 760.
Location/Qualifiers
1. .762
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5107807"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH MGC 12"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."

FEATURES
source
ORIGIN

Query Match          26.8%; Score 705; DB 4; Length 762;
Best Local Similarity 97.6%; Pred. No. 1.1e-172;
Matches 737; Conservative 0; Mismatches 15; Indels 3; Gaps 2;

QY 319 CCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCTTACCTCTGTCTCCCTGG 378
DB 7 CTTGGCTTTGAGCAAGGAGGAGTCTGGATGTGGGAACCTTACCTCTGTCTCCCTGG 66
QY 379 AGCCAGCCCTCCGAGAAACCCAGCTACACGAGCGGAAAGAACCGTCCCGAGCAGG 438
DB 67 AGCCAGCCCTCCGAGAAACCCAGCTACACGAGCGGAAAGAACCGTCCCGAGCAGG 126
QY 439 AACTCAGCAGCTGGAATAGAAATTATTCCTCAACAGCCTGTCACAGCCATTTTCTCTGG 498
DB 127 AACTCAGCAGCTGGAATAGAAATTATTCCTCAACAGCCTGTCACAGCCATTTTCTCTGG 186
```



```
Qy 518 CTCGTGCTCAGCATCGACAGCCTCTCCCTGTGCGATGACAGTGAAGAACCCTCAAA 577
Db 104 CTCGTGCTCAGCATCGACAGCCTCTCCCTGTGCGATGACAGTGAAGAACCCTCAAA 163
Qy 578 GGCCTCTCAAAAGCTCGCGGACACCTCAGCTCAGCGCTACACTCTCGAGCAGCCAGGC 637
Db 164 GGCCTCTCAAAAGCTCGCGGACACCTCAGCTCAGCGCTACACTCTCGAGCAGCCAGGC 223
Qy 638 CGAGCTCGAAGCTCCAGCTGGAACATGCTGTGCCCCGGGGGGCCACCGACACCCC 697
Db 224 CGAGCTCGAAGCTCCAGCTGGAACATGCTGTGCCCCGGGGGGCCACCGACACCCC 283
Qy 698 AAGCTATTTCATGCTGAAGTCCAATACAGTCTCTTAATGGTGAACACCTGCGCAT 757
Db 284 AAGCTATTTCATGCTGAAGTCCAATACAGTCTCTTAATGGTGAACACCTGCGCAT 343
Qy 758 CCGGAGTTCCACCGGGTCAAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCC 817
Db 344 CCGGAGTTCCACCGGGTCAAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCC 403
Qy 818 AGCTGACGCTTCAGCTTGGTCAACAAAGACGGGAGCCTGTTCTGCTACGACATGGAGGT 877
Db 404 AGCTGACGCTTCAGCTTGGTCAACAAAGACGGGAGCCTGTTCTGCTACGACATGGAGGT 463
Qy 878 GCCAGCTCGGGCATCGACTGCGAGTGCACACTGGCCCCCTGATGCGACCTTCGCTGGAG 937
Db 464 GCCAGCTCGGGCATCGACTGCGAGTGCACACTGGCCCCCTGATGCGACCTTCGCTGGAG 523
Qy 938 CTGAGGGTCAAGCATGCGCAGCTGGAGAACAGAGCCCTTAACCTGCCCTCCACCGCGGC 997
Db 524 CTGAGGGTCAAGCATGCGCAGCTGGAGAACAGAGCCCTTAACCTGCCCTCCACCGCGGC 583
Qy 998 TCCACACTGCGGAAAGCAGCCTTCTGCTCGGTGACAGTGTGCTGCCCTGAAACACAGG 1057
Db 584 TCCACACTGCGG--AAGCAGCCTTCTGCTCGGTGACAGTGTGCTGCCCTG--AAACACAGG 641
Qy 1058 CTCAGCGCTTCCAGGGGATGTGNCACGCCGCCCGCTCAGCAGTGGGAACACAGGCGCT 1117
Db 642 CTCAGCGCTTCCAGGG--ATCTGCCAGCCCCCGGCTCAGCAGTGGGAACACAGGCGCTCG 698
Qy 1118 CGNAGCAGNAGTGGGGGCAAGCAGTGGGAGTGGTGGGAGTGGGAGTGGGAGTGGGAGT 1177
Db 699 CAGCAGCAG--GTGGGGGCAAGCAGTGGGAGTGGTGGGAGTGGGAGTGGGAGTGGGAGT 748
Qy 1178 CCNTGCCCCANCCCTGTGTAADAAAACAYTNCGCCCGCAGCTGAAGAGACAGAGGAGATG 1237
Db 749 GCCCTGCCCCACCTGCTGAGAAACACT--CCGCCAGCTGAAGAGACA----GAGGAGG 802
Qy 1238 GNCAGGATTTNACCTTGGGGGAAACAAACAGGGATCTTTTCTGCCCCCTGCTCCAGTN 1297
Db 803 ATGGCAGGAGTTAGCTCGGGGAAACAAACAGG-----ATCTTCTGCCCCCTGCTCAGT 856
Qy 1298 CGAGTTGCGCTGNACCGCTTGGANTCAGTCACCATTTGTTGGGAGCAGAGCAGGAGAGCA 1357
Db 857 CGAGTTGCGCTTGA--CCGCTTGGATCAGTGGCCATTTGCTGGCAGAC--AGGGAGAGCA 912
Qy 1358 GCTTTCAGCCT--GGGTGAGAGGGGTGGGAGAGCCTTTGGCCCCCTCACCTNCCAGGCT 1416
Db 913 GCTTTCAGCCTGGGTCAAGAGGGGTGGGAGAGCCTTTGGCCCCCTCACCTNCCAGGCT 972
Qy 1417 GCTGTGNAGATGTCAAGTGTGTAAAGGNCCTAAGGNCCTAAGGNCCTAAGGNCCTAAGG 1476
Db 973 GCTGTGAA--AGTNTCAAGTGTGTAAAGGNCCTAAGGNCCTAAGGNCCTAAGGNCCTAAGG 1028
Qy 1477 TNCAGCAGGTATGCCCGCCGNTAGGTAAAGGNCCTAAGGNCCTAAGGNCCTAAGGNCCT 1536
Db 1029 TCAGCAGTATTTGCGGGGCCCTTGGTTA---GGGGGGCCCTCTAAGGNCCTTGGCTGGC 1085
Qy 1537 CCTNCACTNCGGCGAGCTCANCCCTTTTGGGTGTAGGGGAAAGAA 1583
Db 1086 CTAACGGGNCAGGTTACCC--CCCTTTTGGGGGGAAGGGGAAAGAA 1130
```

```
RESULT 9
LOCUS AL556130 1133 bp mRNA linear EST 30-MAR-2004
DEFINITION AL556130 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DK011YM14 5-PRIME, mRNA sequence.
ACCESSION AL556130
VERSION AL556130.3 GI:45860849
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1133)
AUTHORS Li,W.B., Gruber,C., Jeesee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31277934.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8426.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DK011BG07QPl&c=8426.r.
FEATURES
Location/Qualifiers
1..1133
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011YM14"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_lib="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
primer="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 24.2%; Score 636.4; DB 1; Length 1133;
Best Local Similarity 84.7%; Pred. No. 1.1e-154;
Matches 946; Conservative 6; Mismatches 115; Indels 50; Gaps 21;
Qy 824 AGCCTTCAGCTTGTGTACCAAGACGGGAGCCTGTTGCTTACGACATGGAGGTGCCAGA 883
Db 1 AGCCTTCAGCTTGTGTACCAAGACGGGAGCCTGTTGCTTACGACATGGAGGTGCCAGA 60
Qy 884 CTCGGGATCGACTGCGAGTGCACACTGGCCCCCTGATGGGAGCTTCGCTGGAGCTGGAG 943
Db 61 CTCGGGATCGACTGCGAGTGCACACTGGCCCCCTGATGGGAGCTTCGCTGGAGCTGGAG 120
Qy 944 GGTCAAGCATGGCAGCTGGAGAACAGGCCCTTAACCTGCCCTCCACCGCGGTGCCACA 1003
Db 121 GGTCAAGCATGGCAGCTGGAGAACAGGCCCTTAACCTGCCCTCCACCGCGGTGCCACA 180
Qy 1004 CTGCCGAAAGCAGCCTTCTGCTGCTGTCACGATGTGCTGCCCTGAAACACAGGCTCAGC 1063
Db 181 CTGCCG--AAGCAGCCTTCTGCTGCTGTCACGATGTGCTGCCCTG--AAACACAGGCTCAGC 238
Qy 1064 CGTTCCAGGGATTTGNCAGCCCCCGGCTCAGCAGTGGGAAACAGGCGCTCGNCAG 1123
Db 239 CGTTCCAGGG--ATCTGCCAGCCCCCGGCTCAGCAGTGGGAAACAGGCGCTCGCAGCAG 295
Qy 1124 CNAGCNAAGTNGGGGCAAGNAGATGCTCTCCAGAGTTTTCACANCTCAGCCNTGC 1183
Db 296 CAAG-----GTGGGGGCAAGCAGNAGATGCTCTCCAGAGTTTTCACAC--TGAGCCCTG 345
Qy 1184 CCANCCCTGCTGAADAAAACAYTNCGCCCGCAGTGAAGAGACAGAGGAGGATGNCAGG 1243
```



```

Db      2678 CTTTCAGCTGGTGCACCAAGATGACAGCCCTGTTGCTATGACATGAGGTGCCAGACT 2737
Qy      886 CGGCATCGACCTGCAGTGCACTGGCCCTGATGGCAGCTTCGCCCTGGAGCTGGAGGG 945
Db      2738 CGGCATCGACCTGCAGTGCACTGGCCCTGATGGCAGCTTTGCTTGGACCTGGAGGG 2797
Qy      946 TCAAGCATGGCCAGCTGGAGACAGGCCCTA 976
Db      2798 TCAAGCATGGTGCAGCTGGAGAACCGACCTA 2828

RESULT 11
LOCUS   CK979910
DEFINITION 4111556 BARC 9BOV Bos taurus cDNA clone 9BOV40_A19 5', mRNA
ACCESSION CK979910
VERSION   CK979910.1 GI:45497890
KEYWORDS EST.
SOURCE    Bos taurus (cow)
ORGANISM  Bos taurus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
          Bovinae; Bos.
REFERENCE 1 (bases 1 to 702)
          Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
          G.P., Bosak, S., Rubenfield, M. and Gasbarre, L.C.
          Production of EST from cDNA libraries derived from immunologically
          activated bovine gut
          Unpublished (2004)
AUTHORS  Contact: Tad S. Sonstegard
          Bovine Functional Genomics Laboratory
          Animal and Natural Resources Institute
          Bldg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA
          Tel: 3015048416
          Fax: 3015048414
          Email: tads@anri.barc.usda.gov
          Single pass sequencing. Bases called and trimmed with phred
          0.000925 using options -trim alt " -trim fasta. Vector identified
          by cross match using options -minmatch 12 -minscore 18
          Plate: 40 row: A column: 19
          Seq primer: CCGATCAGCAGCTGTAAACG
          High quality sequence stop: 702.
FEATURES             source
          1..702
             /organism="Bos taurus"
             /mol_type="mRNA"
             /strain="Holstein"
             /db_xref="taxon:9913"
             /clone="9BOV40_A19"
             /sex="Male"
             /tissue_type="pooled"
             /dev_stage="Multiple"
             /lab_host="DH10B T1 phage resistant"
             /clone_lib="BARC 9BOV"
             /note="Organ: Abomasum; Vector: pAgen-1; Site 1: EcoRV;
             Site 2: NotI; Equimolar amounts of mRNA extracted from
             fundic and pyloric abomasums of 18 and 21 week old steers.
             Exposure to Osteragia osteragi was initiated at 15 weeks
             of age. fundic and pyloric abomasum"

ORIGIN
Query Match      21.6%; Score 568.6; DB 7; Length 702;
Best Local Similarity 88.6%; Pred. No. 5.1e-137;
Matches 616; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy      285 CCACAGAGCCCCCAGAGCAACAAAGTCTCCTCCCTTGACTTTGACGACAGGAGGTCT 344
Db      3 CCTCCAGACCCCCCAGAGCAGACGACAGTCTCCAGCTGCATCGTGGGACAGGAGGTCT 62
Qy      345 GGGATGTGGGACCCCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 404

```

```

Db      63 GGGACATGGGACCACTGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 122
Qy      405 TCACCAGAGCGGAAAGCAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATTA 464
Db      123 TCACCCGAGCGAAAGGCAACCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 182
Qy      465 TTCTCTAACAGCCTGTCTCCAGCCATTTTCTCTCTGAGGAGCAGGAGCAAAATTTCTCTCTGTC 524
Db      183 TTTCTGAACAGCCTGTCTCCAGCCGTTCTCTCTGAGGAGCAGGAGCAGATCTCTCTCTGTC 242
Qy      525 CTGAGCATCGACAGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 584
Db      243 CTGAGCTCGACAGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 302
Qy      585 CAAAGCTCGGGGACACCCCTGAGCTCAGGGGTACACTCTCTGGAGCAGCAGGCGGAGGCT 644
Db      303 CAGAGCTCGGGGACACACTGAGTTCCGGGTGCTATTCTGTGAGCAGCCAGCAGAGGCC 362
Qy      645 CGAAGCTCGAGCTGGAAACATGGTCTGCGCCCGGGGGGGCCACCGACACACCCCAAGCTAT 704
Db      363 CGCAGCTCCAGCTGGAACATGGTCTGCGCCCGGGGGGGCCCACTGACACACCAAGCTAT 422
Qy      705 TTCAATGCTGTGAAGTCCAAATACAGTCTCTTAATGCTGAACACCTGCACATCGGGAG 764
Db      423 TTTCAACGCTGTGAAGTCCAAATACAGTCTCTCAATGGCGAACACCTGCACATCGGGAG 482
Qy      765 TTCCACCGGGTCAAAGTGGGAGACATGCCACTGGCATCAGCAGCCAGATCCCAAGCTCA 824
Db      483 TTCCACCGGGTCAAAGTGGGAGACATGCCAACAGGCATCAGCAGCCAGATCCCAAGCTCA 542
Qy      825 GCCTTCAGCTTGGTCAACAAAGAGCGGAGCCTTTCGCTACGACATGAGAGGTGCCAGAC 884
Db      543 GCCTTCAGCTTGGTCAACAAAGAGCGGAGCCTTTCGCTACGACATGAGAGGTGCCAGAC 602
Qy      885 TCGGGCATCCAGCTGCAGTGCACACTGCGCCCTCATGCGCAGCTTCGCTGAGCTGGAGG 944
Db      603 TCAGGCATCCAGCTGCAGTGCACCTTGGCCCGCCAGCAGCTTCGCTTGGAGCTGGAGG 662
Qy      945 GTCAAGCATGCCAGCTGGAGAACAGCCCTTAACC 979
Db      663 GTCAAGCATGCCAGCTGGAGAACAGCCCTTAAGC 697

RESULT 12
LOCUS   BI161002
DEFINITION 602865208F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5019481 5',
          mRNA sequence.
ACCESSION BI161002
VERSION   BI161002.1 GI:14621003
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 806)
          NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS  National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
          Email: cgabbs-r@mail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Ling Hong/Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: L1CM1834 row: o column: 02
          High quality sequence stop: 600.
          Location/Qualifiers
             1..806
                /organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5019481"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_42"
/notes="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. |"

ORIGIN
Query Match      21.6%; Score 568.2; DB 4; Length 806;
Best Local Similarity 91.9%; Pred. No. 6.8e-137;
Matches 701; Conservative 0; Mismatches 48; Indels 14; Gaps 9;

QY 1 CCCTCTCTCAGCCAGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCGCGTGT 60
DB 42 CCCTCTCAGCCAGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCGCGTGT 101
QY 61 CTCAGCGGAGCTGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGCTCTGAAGA 120
DB 102 CTGAGCGGAGCTGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGCTCTGAAGA 161
QY 121 GCCTTTGGAGGAGATATAAGAACCAAGACATCCAGCCCAATCAAGCAATATACC 180
DB 162 GCCTTTGGAGGAGATATAAGAACCAAGACATCCAGCCCAATCAAGCAATATACC 221
QY 181 ACCAGACCTCTCCATGCCAGCCGAGAGGCTTTGCGCAAGGGGCCAGGGGCCCGCCAG 240
DB 222 ACCAGACCTCTCCATGCCAGCCGAGAGGCTTTGCGCAAGGGGCCAGGGGCCCGCCAG 281
QY 241 CTGAGGAGACACAGGAGAGCCCTTAAGCTCAGCTCTCTC-TCCACACAGAGCCCCA 299
DB 282 CTGAGGAGACACAGGAGAGCCCTTAAGCTCAGCTCTCTC-TCCAGAGAGCCCCA 341
QY 300 GAGCAACCAAGTCTCTCTCTGACCTTTGAGCAAGAGGAGTCTGGGATGTGGGAACCC 359
DB 342 GAGCAACCAAGTCTCTCTCTGACCTTTGAGCAAGAGGAGTCTGGGATGTGGGAAGCC 401
QY 360 TTACTCTGTCTCTCTCTGAGGCA-GCCCTTGCAGAAACCCAGCTCACCAGAGCGGA 418
DB 402 TTACTCTGTCTCTCTCTGAGGCAATGCCCTTCGAGAAAGCCAGCTCACCAGAGCGGA 461
QY 419 AGCAACCTCTCCGAGAGGAAGTGCAGAGCTGGAATAGAAATTAATTCCTCAACAGCCT 478
DB 462 AGCAACCTCTCCGAGAGGAAGTGCAGAGCTGGAATAGAAATTAATTCCTCAACAGCCT 521
QY 479 GTCCAGCAATTTTCTCTGG-AGAGAGGAGCAAAATTTCTCTGTGCTCAGCATCGACA 537
DB 522 GTCCAGCAATTTTCTCTGGAGGAGAGGCAAAATTTCTCTGTGCTCAGCATCGACA 581
QY 538 GC-CTCTCTCTGTGGATGACAGTGAAGAA-CCCATCAAGGCTCTCAAGCTCGCG 595
DB 582 GCACCTCTCTCTGTGGATGACAGTGAAGAAACCCCATTTAGAGCTCTCAAGTTCGCA 641
QY 596 GG---ACACCTGTAGCTCAGGCG---TACACTCTCTGGAGCAGCCAGGCGAGGCTCGAA 648
DB 642 GCGCAACCTTTAGCTCAGGCGGTACAACCTCTCGAGCAACCAAGGCGAGGTTGCA 701
QY 649 GCTCA- GGTGAAACATGGTGTGCGCGGGGGGGCCACCGACACCCCAAGCTATTTC 707
DB 702 GCTCACAGCTGACCTGTGTGCGCGGGGGGAGGCAACGGAACCCCAAGCTATTTC 761
QY 708 AATGTGTGAAGTCCAAATACAG-TCTTTAATGGTGAACAC 749
DB 762 AATGTGTGAAGTCCAAATACAGCCCTTTACTGTGAACAC 804
```

```

RESULT 13
BP267282
LOCUS
DEFINITION
580 bp mRNA linear EST 16-SEP-2004
Sugano cDNA library, Thyroid JTH Homo sapiens cDNA clone
JTH06618, mRNA sequence.
ACCESSION
BP267282.1 GI:52182514
VERSION
BP267282.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 580)
Suzuki Y., Yamashita R., Shiota M., Sakakibara Y., Chiba J.,
Mizushima Sugano J., Nakai K. and Sugano S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1..580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH06618"
/tissue_type="thyroid"
/cell_line="JTH"
/clone_lib="Sugano cDNA library, thyroid JTH"
/notes="thyroid tumor"

ORIGIN
Query Match      21.4%; Score 564.2; DB 5; Length 580;
Best Local Similarity 99.3%; Pred. No. 6.8e-136;
Matches 577; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 461 ATTATTTCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGAGCAGGAGCAATTTCTC 520
DB 1 ATTATTTCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGAGCAGGAGCAATTTCTC 60
QY 521 GTGCTCAGCATCGACAGCCTCTCCCTCGATGACAGTGAAGACCCATCAAGGC 580
DB 61 GTGCTCAGCATCGACAGCCTCTCCCTCGATGACAGTGAAGACCCATCAAGGC 120
QY 581 CTCTCAAGCTCGGGACACCCCTGAGCTCAGGCGTACCTCTGGAGCAGCAGGCGGA 640
DB 121 CTCTCAAGCTCGGGACACCCCTGAGCTCAGGCGTACCTCTGGAGCAGCAGGCGGA 180
QY 641 GGCTCGAAGCTCAGCTGGAACATGGTGTGGCGCGGGCGGCCACCGACACCCCAAG 700
DB 181 GGCTCGAAGCTCAGCTGGAACATGGTGTGGCGCGGGCGGCCACCGACACCCCAAG 240
QY 701 CTATTTCAATGTGTGAAGTCCAATACAGTCTCTTAATGTGAACACCTGCACATCG 760
DB 241 CTATTTCAATGTGTGAAGTCCAATACAGTCTCTTAATGTGAACCGCTGCACATCG 300
QY 761 GGAGTTCCACCGGGTCAAAAGTGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGC 820
DB 301 GGAGTTCCACCGGGTCAAAAGTGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGC 360
QY 821 TGAGCCTTCAGCTTGGTCAACCAAGACGGGAGCTGTTCCTAGCAGATGAGGTGC 880
DB 361 TGAGCCTTCAGCTTGGTCAACCAAGACGGGAGCTGTTCCTAGCAGATGAGGTGC 420
QY 881 AGACTGGGATCGACTGAGTGACACTGGGCGCTGATGGCAGCTTCGCTGGAGCTG 940
DB 421 AGACTGGGATCGACTGAGTGACACTGGGCGCTGATGGCAGCTTCGCTGGAGCTG 480
QY 941 GAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTTAACCTTGCCTCCACCGCGGCTCC 1000
```



```
Db 481 GAGGCTCAAGCATGCGCAGCTGGAGAACAGGCCCTTAACCTGCTCCACCGCGGCTCC 540
Qy 1001 ACACCTGCGGAAAGCAGCCTTCTCTCTCGGTGACGATGCT 1041
Db 541 ACACCTGCGG-AAAGCAGCCTTCTCTCTCGGTGACGATGCT 580

RESULT 14
AL578455/c
LOCUS
DEFINITION AL578455 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
ACCESSION AL578455
VERSION AL578455.3 GI:46257344
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:31316669.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: secre@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8426.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna/s=C50DK011BG07NP1&c=8426.r.
FEATURES
source
1..1095
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011YM14"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_lib="HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 21.0%; Score 553; DB 1; Length 1095;
Best Local Similarity 82.2%; Pred. No. 7e-133;
Matches 899; Conservative 9; Mismatches 129; Indels 57; Gaps 25;
Qy 1458 GNTTCAGTGAGAACACAGCTTCAGCAGTATGCGCGCCGCTAGGTTAANGCGGCGCCT 1517
Db 1038 GGTTCAGTGAGAACACAGCTTCAGCAGTATGCGCGCCGCTAGGTTAANGCGGCGCCT 987
Qy 1518 CTNAAACCCCTTGCTNGGCTNACCTNGGCCAGCTCANCCCTTTTGGGTGTAGGGGA 1577
Db 986 YCTAACCCCTTGCTNGGCTNACCTNGGCCAGCTCANCCCTTTTGGGTGTAGGGGA 931
Qy 1578 AAGNATGCTTGCCTGGGAGGCTGCCCTGGGTAGATACACACACTTTTCAGGTGT 1637
Db 930 AAGAATG-CTGACCTGGGAGGCT-CCCTGGGTAGATACACACACTTTTCAGGTGT 873
Qy 1638 TGCAACACAGCTCCTGAGTGTGACCTCTGTTTCAGCAGGACCAAGAGTGTGAAGT 1697
Db 872 TGCAACACAGCTCCTGAGTGTGACCTCTGTTTCAGCAGGACCAAGAGTGTGAAGT 813
Qy 1698 GAAGTGGTTCTCAGTNNCCCCAGACATGTGCCCTTTTGTGCTGGCTACCACTCTTCCCCA 1757
```

```
Db 812 GAATGGTTCAGT-CCCCAGACATGTGCCCTTTTGTGCTGGCTACCACTTCCCCA 754
Qy 1758 GAGCAGAGGCGCCGAGGCCCTTTCAGGCCAGCACTGCCCCAGA-CTGCTGGCACTCAG 1816
Db 753 GAGCAGAGGCGCCGAGGCCCTTTCAGGCCAGCACTGCCCCAGANCTNGTGGNACTCAG 694
Qy 1817 TTCCCTCATCTGTAAAGGTGAAGGTGATGAGGATATGCTGTGACAGAAACAGTCTGTGG 1876
Db 693 TTCCCTCATCTGTAAAGGTGAAGGTGATGAGGATATGCTGTGACAGAAACAGTCTGTGG 634
Qy 1877 ATGCACATGATCAGTGTCTNAGGNAAGCAGCAGAGAGAGAGCGTCCCGGCCCCAGNCC 1936
Db 633 ATGCACATGATCAGTGTCTAA--GGAAAGCAGCAGAGAGAGAGCGTCCCGGCCCCAGGCC 576
Qy 1937 CCACTNATCAGTGTNCCAGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1996
Db 575 CACTATCAGTGT---CCAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
Qy 1997 CTGACATNACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2056
Db 526 NACTGANACTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 470
Qy 2057 ACNTCTGATGNACCTCAAGCAGCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2116
Db 469 AC-TCTGATG-ACCTCAAGAGMACCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 412
Qy 2117 GGCAGTGCACANCTGTAGGAGCATANGCAAGCCAGAGATGGGTGCTGCTGCTGCTGCTGCTGCTGCT 2176
Db 411 TGACACT-----GTAGGAGCATAGCAGCAGGAGATGGGTGCTGCTGCTGCTGCTGCTGCTGCT 364
Qy 2177 GTCTTGAGCTGTCAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2236
Db 363 GTCTTGAGCTGTCCA-CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
Qy 2237 GAATAGCANCCCTTNNCCCATTTGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2296
Db 308 GAATAGCACCCCTC--TTCCCATTTGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 252
Qy 2297 CAGGACTCAGCTAGCATTAATCAGCTGTGNAATCGTCAGGGGTGCTGCTGCTGCTGCTGCTGCTGCT 2356
Db 251 CAGAACTCAGCTAGCATTAATCAGCTGTG-ATCTGTCAGGGGTGCTGCTGCTGCTGCTGCTGCTGCT 193
Qy 2357 CCTCTGGGCGAGGGGACGCGAGACTCCGTGGGAGAGCTCATTTCCACACATCTTGCCAA 2416
Db 192 CCTCTGGGCGAGGGGACGCGAGACTCCGTGGGAGAGCTCATTTCCACACATCTTGCCAA 133
Qy 2417 GACAGCCTTTTNGTCCAGCTGTCCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2476
Db 132 GACAGCCTTTT-GTCCAGCTGTCCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 74
Qy 2477 CCCCAGCACATAAAGAACTGCGCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2536
Db 73 CCCCAGCACATAAAGAACTGCGCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 15
Qy 2537 TTGTAAGCAATAAA 2550
Db 14 TAAAGNNAARAAA 1

RESULT 15
BU541005
LOCUS
DEFINITION AGENCOURT_10326204 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:6572558
5', mRNA sequence.
ACCESSION BU541005
VERSION BU541005.1 GI:22851446
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 955)
```


Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1899	99.6	1906	6	A66645	A66645 Sequence 1
2	1742	91.4	1782	6	AX089604	AX089604 Sequence
3	1555	81.6	2873	6	AX089606	AX089606 Sequence
4	1555	81.6	3139	6	AX089605	AX089605 Sequence
5	1138.2	59.7	2248	6	BD157657	BD157657 Primer fo
6	1138.2	59.7	2248	6	AX879111	AX879111 Sequence
7	1138.2	59.7	2248	6	AX022425	AX022425 Homo sapi
8	909.4	47.7	1033	6	CQ714292	CQ714292 Sequence
9	794.4	41.7	5833	6	HSW808428	HSW808428 Homo sapi
10	792.8	41.6	6045	6	BD160227	BD160227 Primer fo
11	792.8	41.6	6045	6	AX9883269	AX9883269 Sequence
12	792.8	41.6	6045	9	AK0233827	AK0233827 Homo sapi
13	672.2	35.3	730	6	BD146965	BD146965 Primer fo
14	672.2	35.3	730	6	AX866903	AX866903 Sequence
15	630.8	33.1	1229	9	BC041583	BC041583 Homo sapi
c 16	376	19.7	161864	9	AC007601	AC007601 Homo sapi
c 17	376	19.7	174136	2	AC126766	AC126766 Homo sapi
18	376	19.7	198521	2	AC145723	AC145723 Homo sapi
19	376	19.7	201237	2	AC141081	AC141081 Homo sapi

181 AAACAGTCCAGATCCGCTTTGAGGAGAGAAAGAGATTGGCTCGGATTCGACAGCAGG 240
241 GTCACTGTGTGTGTGTCAGCAATTTGAAGCCCTCTGAGCAATGCTTGAAGAGAGTGA 300
241 GTCACTGTGTGTGTGTCAGCAATTTGAAGCCCTCTGAGCAATGCTTGAAGAGAGTGA 300
301 GGATTTGGACATCACAGCGGCGAGCATCAAGCAGAGCAGCGGGCTTTGGCAGCAAAACCGAA 360
301 GGATTTGGACATCACAGCGGCGAGCATCAAGCAGAGCAGCGGGCTTTGGCAGCAAAACCGAA 360
361 ACAGAGCCCGTGTCTGTGTAAGTCTAGTGAAGAGAGTCTCTCAACAGACAGAGCTCGACGC 420
361 ACAGAGCCCGTGTCTGTGTAAGTCTAGTGAAGAGAGTCTCTCAACAGACAGAGCTCGACGC 420
421 TTCTACTCCCTGGCCCATCGCTCAGACGTGGCCGGGGTGGCGCTGGCTGGCTGT 480
421 TTCTACTCCCTGGCCCATCGCTCAGACGTGGCCGGGGTGGCGCTGGCTGGCTGT 480
481 GCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCTTGAAGAGTCTCTCAACAGACAGAGT 540
481 GCCCTCAACGAACACTCCCTGGAGCGCTACCTGCACATGCTCTTGAAGAGTCTCTCAACAGACAGAGT 540
541 CTGAGCACTTTTATGAAGACTGTCTTTTGTGATGAAGAAAGTCCAGTATGCTT 600
541 CTGAGCACTTTTATGAAGACTGTCTTTTGTGATGAAGAAAGTCCAGTATGCTT 600
601 CCTACCATGGCAGAGTCTGAATCCCATACTCTTTGGATTAACATCGACAAACAGAGT 660
601 CCTACCATGGCAGAGTCTGAATCCCATACTCTTTGGATTAACATCGACAAACAGAGT 660
661 TTGAACGGGCGAGAGTGAAGTCTTCCACCGCTTTCAGACCTCTTAAAGAGTCAACGCAG 720
661 TTGAACGGGCGAGAGTGAAGTCTTCCACCGCTTTCAGACCTCTTAAAGAGTCAACGCAG 720
721 AAGCTGACCTCTTCTGCTCAAGAGTCTTCAAGAGTCTTCAAGAGTCTTCAAGAGTCTTCAAGAGTCT 780
721 AAGCTGACCTCTTCTGCTCAAGAGTCTTCAAGAGTCTTCAAGAGTCTTCAAGAGTCTTCAAGAGTCT 780
781 ACAGCTCTCTGCGCTCTCCATCTCATCAAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTG 840
781 ACAGCTCTCTGCGCTCTCCATCTCATCAAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTG 840
841 TCGTCTCAGGAATCTCAGTGTGATGCAATGCAAAAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
841 TCGTCTCAGGAATCTCAGTGTGATGCAATGCAAAAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
901 AAGTGACCAACATAATCTCATTTGATGATGAGAGAGATGAGAGAACTCTCGGGACGTGT 960
901 AAGTGACCAACATAATCTCATTTGATGATGAGAGAGATGAGAGAACTCTCGGGACGTGT 960
961 TTAAAAAGACACTCGGGCAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCA 1020
961 TTAAAAAGACACTCGGGCAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCA 1020
1021 ATATCATGTCCGCTTTGAAAGCCCTTTGGGCCCTTAATCCAAATGGAATCAGAGCAGAA 1080
1021 ATATCATGTCCGCTTTGAAAGCCCTTTGGGCCCTTAATCCAAATGGAATCAGAGCAGAA 1080
1081 CTATCGAAATATGATTCCTGTCTTTGAAAGAGTCTTGGTACCAAGAGCTTGATGT 1140
1081 CTATCGAAATATGATTCCTGTCTTTGAAAGAGTCTTGGTACCAAGAGCTTGATGT 1140
1141 GAAAGACATCGATGATGAAGATGTGGATGAAACGAAGATGACGTGATGAAACTCATC 1200
1141 GAAAGACATCGATGATGAAGATGTGGATGAAACGAAGATGACGTGATGAAACTCATC 1200
1201 AGGAAGGAGCAGAGGGCCACTCGAGTCCCGAGAGGCACTTGAAGGGAAACCTGT 1260
1201 AGGAAGGAGCAGAGGGCCACTCGAGTCCCGAGAGGCACTTGAAGGGAAACCTGT 1260
1261 CCTCTCCAGATGCAAGCTGGGCTCCGCTGAAGTGTGTCACATGACTCCGACATCCT 1320
1261 CCTCTCCAGATGCAAGCTGGGCTCCGCTGAAGTGTGTCACATGACTCCGACATCCT 1320

QY 1321 CTTCCCTGTCTCAGTGGCTTCCCTACAGCCAGAGATGCCCCCTCGGAAAGCCTTGA 1380
Db 1321 CTTCCCTGTCTCAGTGGCTTCCCTACAGCCAGAGATGCCCCCTCGGAAAGCCTTGA 1380
QY 1381 GAAACGGACAGGACACAGAGACCACTGTTCTCCGGATCTCTGGATCTCGGTACAGTGTGA 1440
Db 1381 GAAACGGACAGGACACAGAGACCACTGTTCTCCGGATCTCTGGATCTCGGTACAGTGTGA 1440
QY 1441 AGCCAGTCTCCAGGACCAAGAGTCTCTGAGCAGCCTGTTACTTCTGCTCAGTGCCA 1500
Db 1441 AGCCAGTCTCCAGGACCAAGAGTCTCTGAGCAGCCTGTTACTTCTGCTCAGTGCCA 1500
QY 1501 GAGTCCATGACAAATAGTGAATCGCCAGGACCACTGTGGCCATGATGAACAGAGAGAT 1560
Db 1501 GAGTCCATGACAAATAGTGAATCGCCAGGACCACTGTGGCCATGATGAACAGAGAGAT 1560
QY 1561 GAGTGGAGGAGAGAACAGATCACTGCGAAAACCTGCTCGAGTGGAGTGGAGCACTCA 1620
Db 1561 GAGTGGAGGAGAGAACAGATCACTGCGAAAACCTGCTCGAGTGGAGTGGAGCACTCA 1620
QY 1621 GCCGCGCTCCGCAAGAGTGGACCTTGAAGAGAGTGGCTGAACAGAGAGAGCGG 1680
Db 1621 GCCGCGCTCCGCAAGAGTGGACCTTGAAGAGAGTGGCTGAACAGAGAGAGCGG 1680
QY 1681 CAGGCGATGAAGTCCAGGCGCTGGCCAGCTATCTTTGCTATTTTGTGAGGAGATCTAA 1740
Db 1681 CAGGCGATGAAGTCCAGGCGCTGGCCAGCTATCTTTGCTATTTTGTGAGGAGATCTAA 1740
QY 1741 CCCCACGTGAGAACCATGTGTTGGAGAAATGGAGGAGAGAGAAATCCACAGTCTCTGA 1800
Db 1741 CCCCACGTGAGAACCATGTGTTGGAGAAATGGAGGAGAGAGAAATCCACAGTCTCTGA 1800
QY 1801 TAGTCTCATTTGAGTCTCTGATCCAGTCTTCTGAAGCTGTGTTCTCTGAGCTTTT 1860
Db 1801 TAGTCTCATTTGAGTCTCTGATCCAGTCTTCTGAAGCTGTGTTCTCTGAGCTTTT 1860
QY 1861 CATGTATGTGAGCAATAAATTTGCTTTTCATTCCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1906
Db 1861 CATGTATGTGAGCAATAAATTTGCTTTTCATTCCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1906

RESULT 2
AX089604
LOCUS Sequence 4 from Patent WO0116314.
DEFINITION AX089604
ACCESSION AX089604
VERSION AX089604.1 GI:13443796
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Wallach,D., Malinin,N., Sinha,I.W. and Leu,S.
TITLE Iren protein, its preparation and use
JOURNAL Patent: WO 016314-A 4 08-MAR-2001;
YEDA RESEARCH AND DEVELOPMENT Co. LTD. (IL)
FEATURES
Location/Qualifiers
source 1..1782
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 91.4%; Score 1742; DB 6; Length 1782;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1775; Conservative 0; Mismatches 4; Indels 3; Gaps 3;
QY 118 ATGAGCGGATCNCNAGACNATGACAAAAGACAAATTTCTGTGGAGCGACTCTCGATGCA 177
Db 1 ATGAGCGGATCAGAGAACAAATGACAAAAGACAAATTTCTGCTGGAGCGACTCTCGATGCA 60

```
QY 178 GTGAAACAGTCCAGATCCGCTTTNGAGGAGAAAGAGATTGCTCGGATTCGACAGC 237
Db 61 GTGAAACAGTCCAGATCCGCTTTNGAGGAGAAAGAGATTGCTCGGATTCGACAGC 120
QY 238 AGGTCACTGTCTGTGTCGCCAGTTTGAAGCCGCTCTGAGCATGCTTTGAAGAGAGT 297
Db 121 AGGTCACTGTCTGTGTCGCCAGTTTGAAGCCGCTCTGAGCATGCTTTGAAGAGAGT 180
QY 298 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGACGCGGCTTTGCGACAAACC 357
Db 181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGACGCGGCTTTGCGACAAACC 240
QY 358 GAAACAGAGCCGTGTTCTGTGTTACTAGTGAAGAGGTCTCTCAACAGACAGAGTGCAG 417
Db 241 GAAACAGAGCCGTGTTCTGTGTTACTAGTGAAGAGGTCTCTCAACAGACAGAGTGCAG 300
QY 418 CGCTTCTACTCCCTGCGCCACATCGCTCAGAGTGGCGCGGGTCCGCGCTGGCTGCG 477
Db 301 CGCTTCTACTCCCTGCGCCACATCGCTCAGAGTGGCGCGGGTCCGCGCTGGCTGCG 360
QY 478 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACTCTGCACATGCTCTGGCGGACCGCTGC 537
Db 361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACTCTGCACATGCTCTGGCGGACCGCTGC 420
QY 538 AGGTGAGCACTTTTATGAAGACTGTCTTTTGTGATGATGAAGAGGTCCAGTATG 597
Db 421 AGGTGAGCACTTTTATGAAGACTGTCTTTTGTGATGATGAAGAGGTCCAGTATG 480
QY 598 CTTCCTACCATGCGAGAGTCTGAATCCATCTCTTCGATTAACATCGACCAACAG 657
Db 481 CTTCCTACCATGCGAGAGTCTGAATCCATCTCTTCGATTAACATCGACCAACAG 540
QY 658 GATTGAAACGGGAGAGTAAGTTGCTCCCAACCGTTTTCAGACCTCTTAAAGAGTCAACG 717
Db 541 GATTGAAACGGGAGAGTAAGTTGCTCCCAACCGTTTTCAGACCTCTTAAAGAGTCAACG 600
QY 718 CAGAACGTGACCTCTTGTGTAAGAGTCAACGCAAGAGTGAAGAGTCTTTCAGGGAG 777
Db 601 CAGAACGTGACCTCTTGTGTAAGAGTCAACGCAAGAGTGAAGAGTCTTTCAGGGAG 660
QY 778 ATCAGACCTCTCTGCGGCTCCATCTCATCAACCTGAAAGAGAGCGA-CCCTTG 836
Db 661 ATCAGACCTCTCTGCGGCTCCATCTCATCAACCTGAAAGAGAGCGA-CCCTTGT 720
QY 837 CCTGTCTGTCCAGAAATGTCAGTGTGATGCCAAATGCAAAAGAGCGGAAGAGAAA 896
Db 721 CCTGTCTGTCCAGAAATGTCAGTGTGATGCCAAATGCAAAAGAGCGGAAGAGAAA 780
QY 897 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 956
Db 781 AAGAAAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840
QY 957 GTGTTTAAAGACACCTGGGCGAGGAGCTCAGAGCAACTCCGACCGCTCTCT 1016
Db 841 GTGTTTAAAGACACCTGGGCGAGGAGAGCTCAGAGCAACTCCGACCGCTCTCT 900
QY 1017 GTCAATATCATGTCCGCTTTGAAAGCCCTTGGGCTTAACCTCAATGGAA-TCAGAGC 1075
Db 901 GTCAATATCATGTCCGCTTTGAAAGCCCTTGGGCTTAACCTCAATGGAAAGTCAGAGC 960
QY 1076 AGCAACTCATGGAAAATGATTCCCTCTTTGAAACGGGAGTTTGGGTACCAAGAGCTT 1135
Db 961 AGCAACTCATGGAAAATGATTCCCTCTTTGAAACGGGAGTTTGGGTACCAAGAGCTT 1020
QY 1136 GATGTGAAAAGCATCGATGATGAAGATGTGATGAAAACCAAGATGACGTGTATGGAAC 1195
Db 1021 GATGTGAAAAGCATCGATGATGAAGATGTGATGAAAACCAAGATGACGTGTATGGAAC 1080
QY 1196 TCATCAGGAAGGACACAGGGCCACTCCGAGTTCGCCAGAGCCACTCGGAAGGAAAC 1255
Db 1081 TCATCAGGAAGGACACAGGGCCACTCCGAGTTCGCCAGAGCCACTCGGAAGGAAAC 1140
QY 1256 ACCTGCCTCTCCAGATGCACAGCTGGGCTCCGCTGAAGGTGCTGCACAATGACTCCGAC 1315
```

```
Db 1141 ACCTGCCTCTCCAGATGCACAGCTGGCTCCGCTGAAGGTGCTGCAATGACTCCGAC 1200
QY 1316 ATCTCTTCTCCTGTAGTGGCTGCTCTACAGCCAGCAGATGCCCCCTCGGAAGC 1375
Db 1201 ATCTCTTCTCCTGTAGTGGCTGCTCTACAGCCAGCAGATGCCCCCTCGGAAGC 1260
QY 1376 CTGGAAACGGGACAGGACAGAGGACCACTTCTCCGATCCTCGGACTTCGGTACAGT 1435
Db 1261 CTGGAAACGGGACAGGACAGAGGACCACTTCTCCGATCCTCGGACTTCGGTACAGT 1320
QY 1436 GTGGAAGCAGCTCTCCAGCCACGGAAGTCTCTGAGCAGCCTGTTA-CTTCTGCTCA 1494
Db 1321 GTGGAAGCAGCTCTCCAGCCACGGAAGTCTCTGAGCAGCCTGTTA-CTTCTGCTCA 1380
QY 1495 GTGCCAGAGTCCATGATCAATTAGTGAACCTGCGCAGGCCACTGTGGCCATGATGAACAGG 1554
Db 1381 GTGCCAGAGTCCATGATCAATTAGTGAACCTGCGCAGGCCACTGTGGCCATGATGAACAGG 1440
QY 1555 AAGGATGAGCTGGAGGAGGAGAAACAGATCACTCGAAACCTGTCGACGCTGAGATGAG 1614
Db 1441 AAGGATGAGCTGGAGGAGGAGAAACAGATCACTCGAAACCTGTCGACGCTGAGATGAG 1500
QY 1615 CACTCAGCCGCTCCGCGCAGAGGTGGACACCTTGAAGAGAGGTGGCTGAACAGAG 1674
Db 1501 CACTCAGCCGCTCCGCGCAGAGGTGGACACCTTGAAGAGAGGTGGCTGAACAGAG 1560
QY 1675 GAGCGCAGGCGCATGAAGTCCAGGCGCTGCGCAGCTATCTTTGCTATTTTGTGAGGAGA 1734
Db 1561 GAGCGCAGGCGCATGAAGTCCAGGCGCTGCGCAGCTATCTTTGCTATTTTGTGAGGAGA 1620
QY 1735 TTCTAACCCACGTGAGAACCATGTGTGGAGAAATGGAGGAGAGAGAAATCCAAAGT 1794
Db 1621 TTCTAACCCACGTGAGAACCATGTGTGGAGAAATGGAGGAGAGAGAAATCCAAAGT 1680
QY 1795 TCCTGATGATCTCATTTGAGCTCTCTGATCCAGTCTTTCTGAGCTGTGTTCTCTG 1854
Db 1681 TCCTGATGATCTCATTTGAGCTCTCTGATCCAGTCTTTCTGAGCTGTGTTCTCTG 1740
QY 1855 ACTTTTTCATGATGTCAGCCAAATAAATGCTTTTCATTTCCCTTG 1896
Db 1741 ACTTTTTCATGATGTCAGCCAAATAAATGCTTTTCATTTCCCTTG 1782
```

RESULT 3
AX089606
LOCUS AX089606 2873 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 6 from Patent WO0116314.
ACCESSION AX089606
VERSION AX089606.1 GI:13443798
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Wallach,D., Malinin,N., Sinha,I.W. and Leu,S.
TITLE Iren protein, its preparation and use
JOURNAL Patent: WO 0116314-A 6 08-MAR-2001;
YEDA RESEARCH AND DEVELOPMENT CO. LTD. (IL)
FEATURES
source 1. .2873
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 81.6%; Score 1555; DB 6; Length 2873;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 4; Indels 3; Gaps 3;
QY 118 ATGAGCGGATCNCGAAACNATGACAAAGAAATTTCTGCTGGAGCGACTGCTGGATGCA 177

Qy 298 CGAGATTGGCACTCAGCGGCGAGCGATCAAGCAGGCGGGCTTTGGCAGCAAAACC 357
 Db 181 CGAGGATGGCACTCAGCGGCGAGCGATCAAGCAGGCGGGCTTTGGCAGCAAAACC 240
 Qy 358 GAAACAGAGCCGCTGTCTGTACTAGTGAAGAGGTCTCTCAACAGACAGAGCTGCAG 417
 Db 241 GAAACAGAGCCGCTGTCTGTACTAGTGAAGAGGTCTCTCAACAGACAGAGCTGCAG 300
 Qy 418 CGCTTCTACTCCCTGCGCCACATCGCTCTCAGAGCTGGGCGGGTCCGCTTGGCTGCGC 477
 Db 301 CGCTTCTACTCCCTGCGCCACATCGCTCTCAGAGCTGGGCGGGTCCGCTTGGCTGCGC 360
 Qy 478 TGTGCGCTCAACGAACTCCCTGGAGGCTACTCTGACATGCTCTGCGGCGAGCCGCTGC 537
 Db 361 TGTGCGCTCAACGAACTCCCTGGAGGCTACTCTGACATGCTCTGCGGCGAGCCGCTGC 420
 Qy 538 AGGCTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCCAGTATG 597
 Db 421 AGGCTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCCAGTATG 480
 Qy 598 CTTCTACCATGGCAGAGTCTGAATCCATCTCTTTCGATTAACATCGAACCAAG 657
 Db 481 CTTCTACCATGGCAGAGTCTGAATCCATCTCTTTCGATTAACATCGAACCAAG 540
 Qy 658 GATTTGAACGGCAGAGTAAGTTTGTCTCCACCGTTTCAGACCTCTTAAAGGAGTCAAG 717
 Db 541 GATTTGAACGGCAGAGTAAGTTTGTCTCCACCGTTTCAGACCTCTTAAAGGAGTCAAG 600
 Qy 718 CAGAACTGACCTCTTGTGAGGAGTCCACGCAAGGAGTGAAGCAGCTGTTTCAGGGAG 777
 Db 601 CAGAACTGACCTCTTGTGAGGAGTCCACGCAAGGAGTGAAGCAGCTGTTTCAGGGAG 660
 Qy 778 ATCAGACCTCTCTGCGCTCTCCATCTCTCAAACTGAAACAGGAGACCGA-CCCTTG 836
 Db 661 ATCAGACCTCTCTGCGCTCTCCATCTCTCAAACTGAAACAGGAGACCGA-CCCTTG 720
 Qy 837 CTTGTCTGTCCAGGAATGTCACTGTCTGTGATGATGCAAAAGGAGCGGAGAGAAA 896
 Db 721 CTTGTCTGTCCAGGAATGTCACTGTCTGTGATGATGCAAAAGGAGCGGAGAGAAA 780
 Qy 897 AAGAACTGACCAATTAATCTTATTTGATGATGAGGAAGTGAAGCAGAGTCTGCGGAC 956
 Db 781 AAGAACTGACCAATTAATCTTATTTGATGATGAGGAAGTGAAGCAGAGTCTGCGGAC 840
 Qy 957 GTGTTTAAAGACACCTGGGCGCAGGAGAGTCTCAGAGCAACTCCGACCGCTCTCT 1016
 Db 841 GTGTTTAAAGACACCTGGGCGCAGGAGAGTCTCAGAGCAACTCCGACCGCTCTCT 900
 Qy 1017 GTCAATATCATGTCCGCTTTGAAAGCCCTTTCGGGCTTAATCTCCAATGAA-TCAGAG 1075
 Db 901 GTCAATATCATGTCCGCTTTGAAAGCCCTTTCGGGCTTAATCTCCAATGAA-TCAGAG 960
 Qy 1076 AGCAACTCATGGAATAATGATTCCTCTCTTTGACCGGGAGTTTGGTACCGAGACTT 1135
 Db 961 AGCAACTCATGGAATAATGATTCCTCTCTTTGACCGGGAGTTTGGTACCGAGACTT 1020
 Qy 1136 GATGTGAAAGACATCGATGATGAAGATGTGATGAAAGCAAGATGATGATGAAAC 1195
 Db 1021 GATGTGAAAGACATCGATGATGAAGATGTGATGAAAGCAAGATGATGATGAAAC 1080
 Qy 1196 TCATCAGGAAGGAGCAAGGGGCACTCGAGTTCGCGCCGAGAGCCACTGGAAGGAAAC 1255
 Db 1081 TCATCAGGAAGGAGCAAGGGGCACTCGAGTTCGCGCCGAGAGCCACTGGAAGGAAAC 1140
 Qy 1256 ACCTGCTCTCCAGATGCAAGTGGCTCCGCTGAGGTGCTGCAATGATCTCCGAC 1315
 Db 1141 ACCTGCTCTCCAGATGCAAGTGGCTCCGCTGAGGTGCTGCAATGATCTCCGAC 1200
 Qy 1316 ATCTCTTCTCTGTCACTGGCTGGCTCTTACAGCCAGCAGATGCCCCCTCGGAAGC 1375
 Db 1201 ATCTCTTCTCTGTCACTGGCTGGCTCTTACAGCCAGCAGATGCCCCCTCGGAAGC 1260
 Qy 1376 CTGGAGAAACGGGACAGGACAGAGGACCACTTCTCCCGATCTCGGACTTCGGTACAGT 1435

Db 1261 CTGGAGAAAGGAGCAGACAGAGGACCAGTTTCTCCGGATCTTGACTTCCGTACGT 1320
 Qy 1436 GTGGAAGCCAGCTCTCTCAGGCCACGGAAGTCTCTTCTGAGCAGCCTGTTA-CTTCTGCTCA 1494
 Db 1321 GTGGAAGCCAGCTCTCTCAGGCCACGGAAGTCTCTTCTGAGCAGCCTGTTTACCTTCTGCTCA 1380
 Qy 1495 GTGCCAGATTCATGATCAATTAAGTAACTGCGGCGAGCCACTGTGGCATGATGAACAGG 1554
 Db 1381 GTGCCAGATTCATGATCAATTAAGTAACTGCGGCGAGCCACTGTGGCATGATGAACAGG 1440
 Qy 1555 AAGGATGAGCTGGAGGAGGAGAACAGATCACTCGGAAACCTGTGACAGGTGAGATGGAG 1614
 Db 1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTCGGAAACCTGTGACAGGTGAGATGGAG 1500
 Qy 1615 CACTCAGCCGCTCCGCAAGAGGTGGACACCTTGAAGAGGAGTGGCTGAACAGGAG 1674
 Db 1501 CACTCAGCCGCTCCGCAAGAGGTGGACACCTTGAAGAGGAGTGGCTGAACAGGAG 1560
 Qy 1675 GAGCGGAGGATGAAGGTCCAGGCGCTGGCCAG 1709
 Db 1561 GAGCGGAGGATGAAGGTCCAGGCGCTGGCCAG 1595

RESULT 5
 BD157657
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BD157657 2248 bp DNA linear PAT 17-JAN-2003
 Primer for synthesizing full-length cDNA and use thereof.
 BD157657
 BD157657.1 GI:27863415
 JP 2002191363-A/12500.
 Homo sapiens (human)
 Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

1 (bases 1 to 2248)
 Oca, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
 Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 12500 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 FN JP 2002191363-A/12500
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORI
 SAITO.
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 KEIICHI NAGAI, TETSUJI OTSUKI
 PC
 C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10,
 C12N5/09, C12P21/02, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof PH Key

FT CDS
 Location/Qualifiers
 (63)..(1187).
 1..2248
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

FEATURES
 source
 1..2248
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 59.7%; Score 1138.2; DB 6; Length 2248;
 Best Local Similarity 94.7%; Pred. No. 5.4e-271;
 Matches 1195; Conservative 0; Mismatches 65; Indels 2; Gaps 2;
 Qy 77 GCGCGGCGCGCGCAGCAGCCCGCGCGGAGAGCCACCATGAGCGGATCAGAGAACN 136
 Db 22 GCGCGGCGCGCGCAGCAGCCCGCGCGGAGAGCCACCATGAGCGGATCAGAGAAC 81
 Qy 137 ATGACAAAAGACAATTTCTGCTGAGGACTGCTGGATGAGTGAACAGTGCAGATCC 196

```
Db 82 ATGCAAAAAGACAAATTCCTGCTGGAGCGACTGCTGGATGCACTGAAACAGTGCAGATCC 141
Qy 197 GCTTTTNGAGGAGAAAGAGATTGCTCGGATTCCGACAGAGGTCACCTGTCTGTGTG 256
Db 142 GCTTTGAGGAGGAGAAAGAGATTGCTCGGATTCCGACAGAGGTCACCTGTCTGTGTG 201
Qy 257 CCCAGTTTGAAGCCGTCCTGACGATGGCTTGAAGAGGAGTCGAGATTGGCACTCACAG 316
Db 202 CCCAGTTTGAAGCCGTCCTGACGATGGCTTGAAGAGGAGTCGAGATTGGCACTCACAG 261
Qy 317 CGCAGCGATCAAGCAGGACGGGCTTTGCGACAAAACCGAAACAGAGCCCGTGTCT 376
Db 262 CGCAGCGATCAAGCAGGACGGGCTTTGCGACAAAACCGAAACAGAGCCCGTGTCT 321
Qy 377 GGTACTAGCTGAAGAGTCTCTACACAGCAGCAGCTCAGGCTTCTACTCCCTGCGCC 436
Db 322 GGTACTAGCTGAAGAGTCTCTACACAGCAGCAGCTCAGGCTTCTACTCCCTGCGCC 381
Qy 437 ACATCGCTCAGACGCTGGGCGGGGTCGGCTGCTGCTGCTGCTGCTCAACGAACACT 496
Db 382 ACATCGCTCAGACGCTGGGCGGGGTCGGCTGCTGCTGCTGCTCAACGAACACT 441
Qy 497 CCCTGGAGGCTACTGACATGCTCTCTGGCGGACCGCTGAGGCTGAGCACTTTTATG 556
Db 442 CCCTGGAGGCTACTGACATGCTCTCTGGCGGACCGCTGAGGCTGAGCACTTTTATG 501
Qy 557 AAGACTGCTCTTTGTGATGATGAAGAAAGTCCAGTATGCTTCTTACCATGCGCAG 616
Db 502 AAGACTGCTCTTTGTGATGATGAAGAAAGTCCAGTATGCTTCTTACCATGCGCAG 561
Qy 617 GTCTGAACCTCACTACTCTTTGCGATTACATCGCAACAAAGATTGTAAACGGGACAGTA 676
Db 562 GTCTGAACCTCACTACTCTTTGCGATTACATCGCAACAAAGATTGTGAACGGGACAGTA 621
Qy 677 AGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGAGAACTGACCTCTTGC 736
Db 622 AGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGAGAACTGACCTCTTGC 681
Qy 737 TGAAGGAGTCACGCAAGGAGTGAAGCAGCTGTTTCAAGGAGTCAACGAGAACTGACCTCTTGC 796
Db 682 TGAAGGAGTCACGCAAGGAGTGAAGCAGCTGTTTCAAGGAGTCAACGAGAACTGACCTCTTGC 741
Qy 797 TCTCATCTCTCATCAAACTGAACAGGAGACCGA - CCCTTGCCTGTGTGTCTCAGGAATG 855
Db 742 TCTCATCTCTCATCAAACTGAACAGGAGACCGA - CCCTTGCCTGTGTGTCTCAGGAATG 801
Qy 856 TCAGTGTCTGATGCCAAATGCAAAAGGAGCGGAGAGAGAAAGAGTGACCAACATAA 915
Db 802 TCAGTGTCTGATGCCAAATGCAAAAGGAGCGGAGAGAGAGAAAGTGACCAACATAA 861
Qy 916 TCTCATTTGATGATGAGGAAGATGAGCAGAACTCTTGGGAGCTGTTTAAAGACACCTG 975
Db 862 TCTCATTTGATGATGAGGAAGATGAGCAGAACTCTTGGGAGCTGTTTAAAGACACCTG 921
Qy 976 GGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAATATCATGTCCCGCT 1035
Db 922 GGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAATATCATGTCCCGCT 981
Qy 1036 TTGAAGCCCTTTCGGGCTAACTCCAAATGAA - TCAGAGCAGCAACTCATATGGAATTTG 1094
Db 982 TTGAAGCCCTTTCGGGCTAACTCCAAATGAAAGTCAAGAGCTGATGTAAGAGCATCATG 1154
Qy 1095 ATTCCCTGTCTTTGAAACGGGAGTTTGGGTACAGAGCTTGTGTAAGAGCATCATG 1154
Db 1042 ATTCCCTGTCTTTGAAACGGGAGTTTGGGTACAGAGCTTGTGTAAGAGCATCATG 1101
Qy 1155 ATGAAGATGTGGATCAAAACGAAGATCAGCTGTATGGAATCTCATGGAAGGAGGAGCA 1214
Db 1102 ATGAAGATGTGGATCAAAACGAAGATCAGCTGTATGGAATCTCATGGAAGGAGGAGCA 1161
Qy 1215 GGGGCCACTCGGAGTCGCGCAGAGCACTGGAAGGGAACCACTGCTCTCCAGATGC 1274
Db 1162 GGGGCCACTCGGAGTCGCGCAGAGTAAAGTTTGTGTGTAAGGTTGAGTCTCAACCGTCCC 1221
```

```
Qy 1275 ACAGCTGGCTCCCTGAGGTGCTGCACATGATGATCGGACATCTCTTCCCTGTCAGTG 1334
Db 1222 CCAGGCTGGAGTGCCTGGCTGATCTCAGCTCACTGCACTCCACCTCCCGGTTTCAG 1281
Qy 1335 GC 1336
Db 1282 GC 1283

RESULT 6
AX879111
LOCUS AX879111 2248 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 14016 from Patent EP1074617.
ACCESSION AX879111
VERSION AX879111.1 GI:40033847
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Carnivora; Homini; Homo.
1
REFERENCE
AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primers for synthesising full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 14016 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source Location/Qualifiers
1..2248
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
63..1190
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAE90370.1"
/db_xref="GI:40033848"
/translation="MSGSONDKRQELERLLDAVKOCQIRFGGRKEIASDSRVTC
LCAQFVAVLHGLKRSRGLATAAAKQAGAFASKTETFPVFTYFVKEVNLKHLQRM
YSLRHASDVGRGRAMLRCALNLSUERYLHMLADRLSLTFFVYFVWDESRSM
LPTMAAGLSILIFAINIDKDLQSKFAPVSDLLKSTQNTVSLKSTQGVSSIF
RETTASAVSILIKPQETDPLPVSRNVSADAKKERRKKKVTNIIIFDDDEDQ
NSGVFKTTPGAGESSDRSVNIMSAFSPFGPNSNGSSQSSNSWKIDSLSLNGE
FGYQKLDVKGIDDEDVDENEDDYVYGNSSGRKGRHSSEPEK"
ORIGIN
Query Match 59.7%; Score 1138.2; DB 6; Length 2248;
Best Local Similarity 94.7%; Pred. No. 5.4e-271;
Matches 1195; Conservative 0; Mismatches 65; Indels 2; Gaps 2;
Qy 77 GCCGGCGCGCGCGGCGGACCGCGCCCGGGAGAGCCATGAGCGGATCNGAACN 136
Db 22 GCGGCGCGCGCGCGGCGGACCGCGCCCGGGAGAGCCATGAGCGGATCAGAGAA 81
Qy 137 ATGACAAAAGACAAATTTCTCTGAGCGCATCTGCTGAGTGCAGTGAACAGTCCAGATCC 196
Db 82 ATGACAAAAGACAAATTTCTCTGAGCGCATCTGCTGAGTGCAGTGAACAGTCCAGATCC 141
Qy 197 GCTTTNGAGGAGGAGGAGGAGATTGCTTCCGATCCGACAGAGGTCACCTGTCTGTGTG 256
Db 142 GCTTTGAGGAGGAGGAGGAGGATTGCTTCCGATCCGACAGAGGTCACCTGTCTGTGTG 201
Qy 257 CCAGATTGAGCGCTCTCGACGATGGCTTGAAGAGGATCGAGGATCGGACTTCACAG 316
Db 202 CCAGATTGAGCGCTCTCGACGATGGCTTGAAGAGGATCGAGGATCGGACTTCACAG 261
Qy 317 CGCAGCGATCAAGCAGGACGGGCTTTGCCAGCAAAACCGAAACAGAGCCCGTGTCT 376
Db 262 CGCAGCGATCAAGCAGGACGGGCTTTGCCAGCAAAACCGAAACAGAGCCCGTGTCT 321
Qy 377 GGTACTAGTGAAGGAGGTCCTCAACGAAGCAGAGCTGCGAGCGCTTCTACTCTCCGCGCC 436
Db 322 GGTACTAGTGAAGGAGGTCCTCAACGAAGCAGAGCTGCGAGCGCTTCTACTCTCCGCGCC 381
```


/protein_id="BAB14033.1"
/db_xref="GI:10433819"
/translation="MSGSONDKROFLERLLLDVAKQCIIRFGGRKEIASDSRVTCLCAQFEAVLHGLGKSRGLATAAIAKQAGPASKTETEPVPVYVYKVLNKHLORFYSLRHIASDVGGRALWLCALNEHSLERYLHMLLADRLSTFFVDFWVDEERSMLPTMAAGLNSILFPAINDKIDJGSGKFAPTVSDLLKESQNTVSLKESQGVSSSLFREITASSAVSILIKPEQETDPLPVVSRVNSADAKCKKRRKKKVTNIISFDDDEDEQNSGDFKKTTPGAGESSENSDRSVNIMSAPFSPGPNSSNGSQSNKWKIDLSLNGEFGYKLDVKGIDDEDVEDNEDVDVYSSGRKXHRGHSPEK"
ORIGIN

Query Match 59.7%; Score 1138.2; DB 9; Length 2248;
Best Local Similarity 94.7%; Pred. No. 5.4e-271;
Matches 1195; Conservative 0; Mismatches 65; Indels 2; Gaps 2;
QY 77 GCGCGCGCGCGCGCANGCACCGCCGCGGANNAGCNCCATGAGCGGATCNGNACN 136
DB 22 GCGCGCGCGCGCGCAGCGCCGCGGAGAGGACCATTGAGCGGATCAGAAACA 81
QY 137 ATGCAAAAGACAATTTCTGTGGAGCGACTGCTGGATGCAATGCAAGTGCAGATCC 196
DB 82 ATGCAAAAGACAATTTCTGTGGAGCGACTGCTGGATGCAATGCAAGTGCAGATCC 141
QY 197 GCTTTNGAGGAGAAAGAGATTGCTCGGATTCGACAGAGGTCACCTGTCTGTGTG 256
DB 142 GCTTTGAGGAGGAGAAAGAGATTGCTCGGATTCGACAGAGGTCACCTGTCTGTGTG 201
QY 257 CCCAGTTTGAAGCCGCTCTGCAGCATGCTTGAAGAGGAGTCGAGATTGGCACTCACAG 316
DB 202 CCCAGTTTGAAGCCGCTCTGCAGCATGCTTGAAGAGGAGTCGAGATTGGCACTCACAG 261
QY 317 CGCGAGCGATCAAGCAGCAGCGGCTTTGCCAGCAAAACCAACAGAGCCCGTGTCT 376
DB 262 CGCGAGCGATCAAGCAGCAGCGGCTTTGCCAGCAAAACCAACAGAGCCCGTGTCT 321
QY 377 GGTACTAGCTGAGAGAGTCTCTACACAGCAGAGCTGCAGCGCTTCTACTCCCTGGGCC 436
DB 322 GGTACTAGCTGAGAGAGTCTCTACACAGCAGAGCTGCAGCGCTTCTACTCCCTGGGCC 381
QY 437 ACATCGCTCAGACGTGGCGCGGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496
DB 382 ACATCGCTCAGACGTGGCGCGGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
QY 497 CCCTGGAGCGTACTGTCACATGCTCTGCGCGCAGCGCTGCAGCGTGCAGCTTTTATG 556
DB 442 CCCTGGAGCGTACTGTCACATGCTCTGCGCGCAGCGCTGCAGCGTGCAGCTTTTATG 501
QY 557 AAGACTGTCTTTGTGTGATGATGAAGAAAGTCCAGTATGCTTCTTCTTCTTCTTCTTCTTCT 616
DB 502 AAGACTGTCTTTGTGTGATGATGAAGAAAGTCCAGTATGCTTCTTCTTCTTCTTCTTCTTCT 561
QY 617 GTCTGAATCCATATCTTTTCGGATTAAACATCGACAACAGGATTTGAAACGGGCGAGTA 676
DB 562 GTCTGAATCCATATCTTTTCGGATTAAACATCGACAACAGGATTTGAAACGGGCGAGTA 621
QY 677 AGTTTGCTCCACCGTTTCAGACCTCTTAAAGAGTCAACCGAGACGTCGACCTCTTCG 736
DB 622 AGTTTGCTCCACCGTTTCAGACCTCTTAAAGAGTCAACCGAGACGTCGACCTCTTCG 681
QY 737 TGAAGGAGTCCACGCAAGGAGTACGCGCTGTTTCAGGAGATCAGACGCTCTCTCTGCG 796
DB 682 TGAAGGAGTCCACGCAAGGAGTACGCGCTGTTTCAGGAGATCAGACGCTCTCTCTGCG 741
QY 797 TCTCCATCTCTCATCAAAACCTGAAACAGGAGACCGA - CCCTTTCCTGCTGCTGCTTCCAGGAATG 855
DB 742 TCTCCATCTCTCATCAAAACCTGAAACAGGAGACCGACCCCTTTCCTGCTGCTGCTTCCAGGAATG 801
QY 856 TCAGTGTCTGTCCTCAATGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 915
DB 802 TCAGTGTCTGTCCTCAATGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 861
QY 916 TCTCATTTGATGATGAGGAGATGAGCAGAACTCTGGGACGCTGTTTTAAAAAGACACCTG 975

DB 862 TCTCATTTGATGATGAGGAGATGAGCAGAACTCTGGGAGCGTGTTTAAAAAGACACCTG 921
QY 976 GGGCAGGGGAGAGCTCAGAGGACAACTCCGACCCGCTCTCTGTCAATATCATGTCCGCT 1035
DB 922 GGGCAGGGGAGAGCTCAGAGGACAACTCCGACCCGCTCTCTGTCAATATCATGTCCGCT 981
QY 1036 TTGAAAGCCCTTCGCGCTTAACTCCCAATGGAA - TCAGAGCAGCAACTCATGGAATAATG 1094
DB 982 TTGAAAGCCCTTCGCGCTTAACTCCCAATGGAACTCAGAGCACTCATGGAATAATG 1041
QY 1095 ATTCCCTGTCTTTGAACGGGAGTTTCGGTACCAGAACTTGATGTGAAAGCATCGATG 1154
DB 1042 ATTCCCTGTCTTTGAACGGGAGTTTCGGTACCAGAACTTGATGTGAAAGCATCGATG 1101
QY 1155 ATGAAGATGTGATGAAACGAGATGACGCTGTATGAAACTCATCAGGAGGAGGAGCACA 1214
DB 1102 ATGAAGATGTGATGAAACGAGATGACGCTGTATGAAACTCATCAGGAGGAGGAGCACA 1161
QY 1215 GGGGCCACTCGGAGTCCCGAGAGCCACTGGAAGGGAAACACCTGCTCTCCAGATGC 1274
DB 1162 GGGGCCACTCGGAGTCCCGAGAGAGTAAGTTTGTGTGAAGTGGAGTCTCACCGTCCC 1221
QY 1275 ACAGCTGGCTCCGCTGAAGTGTGCACAATGACTCCGACATCTCTTCCCTGTCTAGTG 1334
DB 1222 CCAGGCTGAGTGCCTGGCGTGTATCTCAGTCACTGCAACCTCCACCTCCCGGTTTCA 1281
QY 1335 GC 1336
DB 1282 GC 1283
RESULT 8
LOCUS CQ714292 1033 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 226 from Patent WO02068579.
ACCESSION CQ714292
VERSION CQ714292.1 GI:42275149
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 226 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1 .1033
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 47.7%; Score 909.4; DB 6; Length 1033;
Best Local Similarity 91.9%; Pred. No. 2.9e-214;
Matches 1026; Conservative 0; Mismatches 5; Indels 86; Gaps 3;
QY 125 GATCNCNGACNATGACAAAGACAATTTCTGCTGAGCGACTGCTGGATGCAAGTCA 184
DB 1 GATCAGACAACATGACAAAGACAATTTCTGCTGAGCGACTGCTGGATGCAAGTCA 60
QY 185 AGTGCCAGATCCGCTTTNGAGGAGAAAGAGATTCCTCGGATTCGACAGCAGGATCA 244
DB 61 AGTGCCAGATCCGCTTTNGAGGAGAAAGAGATTCCTCGGATTCGACAGCAGGATCA 120
QY 245 CCTGTCTGTGCCCCAGTTTGAAGCGCTCTGTCAGCATGGCTTTGAAGAGGAGTCCAGGAT 304
DB 121 CCTGTCTGTGCCCCAGTTTGAAGCGCTCTGTCAGCATGGCTTTGAAGAGGAGTCCAGGAT 180
QY 305 TGGCACTCAGCGGCGAGCGATCAGCAGCGAGCGGGCTTTTGGCCAGCAAAACCGAAACAG 364


```
Db 181 TGGCRCTCACAGCGGAGGATCAAGCAGGAGCGGGCTTTGCCAGCAAAACCGAA--- 237
Qy 365 AGCCCGTGTCTGCTACTAGTGAAGGAGTCTTCAACAGCAGCAGAGCTGCGGCTTCT 424
Db 238 ----- 237
Qy 425 ACTCCCTGCGCCATCGCTCAGACGTGGGCGGGCTGCGGCTGCGCTGCGTGTGCC 484
Db 238 ----- 276
Qy 485 TCAACGAACACTCCCTGGAGGCTACTGACATGCTCCTGGCGGCGGCTGCGGCTGCA 544
Db 277 TCAACGAACACTCCCTGGAGGCTACTGACATGCTCCTGGCGGCGGCTGCGGCTGCA 336
Qy 545 GCACCTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGTCCAGTATGCTTCTTA 604
Db 337 GCACCTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGTCCAGTATGCTTCTTA 396
Qy 605 CCATGGCAGAGTCTGAACCTCCATCTCTTTGCGATTAACATCGACAACAGATTGA 664
Db 397 CCATGGCAGAGTCTGAACCTCCATCTCTTTGCGATTAACATCGACAACAGATTGA 456
Qy 665 ACGGCGAGAGTAAAGTTTGTCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAAGC 724
Db 457 ACGGCGAGAGTAAAGTTTGTCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAAGC 516
Qy 725 TGACCTCTTGTCTGAAGAGTCCACGCAAGAGTGAAGCAGCTGTTTCAAGGAGATCACAG 784
Db 517 TGACCTCTTGTGAAGAGTCCACGCAAGAGTGAAGCAGCTGTTTCAAGGAGATCACAG 576
Qy 785 CTTCTCTGCGCTCTCATCTCAATCAAACTGAACAGGAGACCGA---CCCTTCCGTGTCG 843
Db 577 CTTCTCTGCGCTCTCATCTCAATCAAACTGAACAGGAGACCGA---CCCTTCCGTGTCG 636
Qy 844 TGTCCAGGAATGTCAGTGTGATGTCGCAAAATGCAAAAGGAGCGGAAGAAAAGAAAG 903
Db 637 TGTCCAGGAATGTCAGTGTGATGTCGCAAAATGCAAAAGGAGCGGAAGAAAAGAAAG 696
Qy 904 TGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGAGCTGTTTA 963
Db 697 TGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGAGCTGTTTA 756
Qy 964 AAAGACACCTGGGCGAGGAGAGTCTAGAGGACAACTCCGACCGCTCTCTGTCAATA 1023
Db 757 AAAGACACCTGGGCGAGGAGAGTCTAGAGGACAACTCCGACCGCTCTCTGTCAATA 816
Qy 1024 TCATGTCCGCTTTGAAGCCCTTCCGGCTTAACCTCAATGGAA---TCAGAGCAGCACT 1082
Db 817 TCATGTCCGCTTTGAAGCCCTTCCGGCTTAACCTCAATGGAA---TCAGAGCAGCACT 876
Qy 1083 CATGGAAAATTTGATTCCTCTCTTTGAACGGGAGTTTGGGTACCAAGCTTTGATGTGA 1142
Db 877 CATGGAAAATTTGATTCCTCTCTTTGAACGGGAGTTTGGGTACCAAGCTTTGATGTGA 936
Qy 1143 AAAGCATTCGATGATGAAGATGTGGATGAACGAAAGATGATGATGAAATCTATCAG 1202
Db 937 AAAGCATTCGATGATGAAGATGTGGATGAACGAAAGATGATGATGAAATCTATCAG 996
Qy 1203 GAAGGAAGCAGAGGCGCACTCGAGTCCGCCAGAA 1239
Db 997 GAAGGAAGCAGAGGCGCACTCGAGTCCGCCAGAA 1033
```

RESULT 9

```
HS0808428
LOCUS HSM0808428 5833 bp mRNA linear PRI 30-AUG-2003
DEFINITION Homo sapiens mRNA; cDNA DKFp686E09125 (from clone DKFp686E09125).
ACCESSION BX648280
VERSION BX648280.1 GI:34367439
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
```

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 5833)
Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Pobo,G., Han,M. and Wiemann,S.
The German Human cDNA Consortium
Direct Submission
Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFp686E09125) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

Location/Qualifiers
source
1..5833
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFp686E09125"
/tissue_type="human uterus"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host DH10B; sites SfiIIA + SfiIB"
/dev_stage="adult"
polya_signal
5789..5794
polya_site
5810

ORIGIN

Query Match 41.7%; Score 794.4; DB 9; Length 5833;
Best Local Similarity 96.9%; Pred. No. 1e-185;
Matches 853; Conservative 0; Mismatches 21; Indels 6; Gaps 4;
Qy 363 AGAGCCCGTGTCTTGTGTTACTAGCTGAAGGAGGTCCTCAACAGCAGCAGCTGCGCTG 421
Db 1381 AGAGCCCGTGTCTTGTGTTACTAGCTGAAGGAGGTCCTCAACAGCAGCAGCTGCGCT 1440
Qy 422 TCTACTCCCTCGCCACATCGCTCAGACGTGGGCGGGGTCGGCTGCGCTGCTG 481
Db 1441 TCTACTCCCTCGCCACATTCGCTCAGACGTGGGCGGGGTCGGCTGCGCTGCTG 1500
Qy 482 CCCTCAACGAACTCCCTGGAGCGCTACTGTGACATGTCTCTGGCGGACCGCTGCAAGC 541
Db 1501 CCCTCAACGAACTCCCTGGAGCGCTACTGTGACATGTCTCTGGCGGACCGCTGCAAGC 1560
Qy 542 TGAGCACTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGTCCAGTATGCTTC 601
Db 1561 TCAGTACTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGAGTCCAGTATGCTTC 1620
Qy 602 CTACCATGGCAGCAGTCTGAACCTCCATCTCTTTGGGATTAACATCGCAACAGGATT 661
Db 1621 CTACCATGGCAGCAGTCTGAACCTCCATCTCTTTGGGATTAACATCGCAACAGGATT 1680
Qy 662 TGAACGGGCGAGAGTAAAGTTTGTCTCCACCGTTTTCAGACCTCTTTAAAGGAGTCAACGAGA 721
Db 1681 TGAACGGGCGAGAGTAAAGTTTGTCTCCACCGTTTTCAGACCTCTTTAAAGGAGTCAACGAGA 1740
Qy 722 AGCTGACCTCTTGTCTGAAGGAGTCCACGCAAGGAGTGAAGCAGCTGTTTCAGGAGATCA 781
Db 1741 ATGTGA---CCTTGTCTGAAGGAGTCCACGCAAGGAGTGAAGCAGCTGTTTCAGGAGATCA 1797
Qy 782 CAGCTCTCTGCGCTCTCCATCTCATCAACCTGAAACAGGAGACCGA---CCCTTGGCTG 840
Db 1798 CAGCTCTCTGCGCTCTCCATCTCATCAACCTGAAACAGGAGACCGACCCCTTGGCCG 1857
Qy 841 TCGTGTCCAGGAGTGTGAGTGTGATGCGCAAAATGCAAAAGGAGCGGAAGAAAAGA 900
Db 1858 TCGTGTCCAGGAGTGTGAGTGTGATGCGCAAAATGCAAAAGGAGCGGAAGAAAAGC 1917

FEATURES source	Location/Qualifiers		ORIGIN
	1..6045	6045 bp mRNA linear	
	/organism="Homo sapiens"	Homo sapiens cdna FLJ13765 fis, clone PLACE4000128, weakly similar to Mus musculus putative transcription factor mRNA.	
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
CDS	364..957		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAE91902.1"		
	/db_xref="GI:40038171"		
	/translations="MRVHCAIIPSDMLHSTNCRTABEMVLTLLDYLFRHEVOAVSSL SGQKHGKQLDPLTYIGIRCHLFYKFGITSDMYRIKOSIDSKCTARWRQWGOSL VYKFSRRPNSSSYCPSPWMTPTPPARELLQPLQPLQALHYALANVOVQLHQICE DGVQVPIQGLHIAKVPGEQVQITQDSSEPCSGTT"		
	Query Match 41.6%; Score 792.8; DB 6; Length 6045;		
	Best Local Similarity 96.8%; Pred. No. 2.6e-185;		
	Matches 852; Conservative 0; Mismatches 22; Indels 6; Gaps 4;		
QY	363 AGAGCCCGTGTTCCTGCTACGTGAAGGAGTCTCTCAACAAGCAGCAGCTGC-AGCGCT	AK023827	1709 TGAAGAGCATCGATGATGAAGATGTGGATGAAACCAAGATGACGTGTATGGAATCAT 1768
Db	932 AAAGCCCGTGTTCCTGCTACGTGAAGGAGTCTCTCAACAAGCAGCAGCTGC	AK023827	1200 CAGGAGGAAGCAGCAGGCGCCACTCGAGTCGCCCGAGAA 1239
QY	422 TCTACTCCCTCGCCACATCGCTCAGACGTGGCGCGGGTCCGCTCGCTGCTGTG	AK023827.1	1769 CAGGACGGAAGCAGCAGGCGCCACTCAGATCGCCCGAGAA 1808
Db	992 TCTACTCCCTCGCCACATCGCTCAGACGTGGCGCGGGCCGACCTGGCTGCTG		
QY	482 CCCTCAACGAACACTCCCTGGAGCGTACTCTGACATGCTCTCTGGCGGCGGCGG		
Db	1052 CCTCAACGAACACTCCCTGGAGCGTACTCTGACATGCTCTCTGGCGGCGGCGG		
QY	542 TGAGCACTTTTATGAAGACTGCTCTTTGTGTGATGATGAAGAAAGTCCAGTATGCTTC		
Db	1112 TCAGTACTTTTATGAAGACTGCTCTTTGTGTGATGATGAAGAGTCCAGTATGCTTC		
QY	602 CTACCATGGCAGAGTCTGAATCTCAATCTTTGGGATTAACATCGAACAAAGATT		
Db	1172 CTACCATGGCAGAGTCTGAATCTCAATCTTTGGGATTAACATCGAACAAAGATT		
QY	662 TGAACGGGCAGAGTAAAGTTGCTCCACCGTTTCAGACCTCTTAAAGGAGTCAACCGA		
Db	1232 TGAACGGGCAGAGTAAAGTTGCTCCACCGTTTCAGACCTCTTAAAGGAGTCAACCGA		
QY	722 ACCTGACCTCTTGTCTGAAGAGTCCACGCAAGAGTGAAGCAGCTGTTCAAGGAGATCA		
Db	1292 ATGTGA---CCTTGCTGAAGAGTCCACGCAAGAGTGAAGCAGCTGTTCAAGGAGATCA		
QY	782 CAGCCTCTCTGCGCTCCATCTCATCAAACTGAACAGGAGACCGA-CCCTTGCCTG		
Db	1349 CAGCCTCTCTGCGCATCTCCATCTCATCAAACTGAACAGGAGACCGACCCCTTGCCTG		
QY	841 TCGTGTCCAGGATGTCAGTGTGATGTCGCAATGCAAAAAGGAGCGGAAGAAAAAGA		
Db	1409 TCGTGTCCAGGATGTCAGTGTGATGTCGCAATGCAAAAAGGAGCGGAAGAAAAAGC		
QY	901 AAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGAGCTGT		
Db	1469 AAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGAGCATGT		
QY	961 TTAAGAAGACACCTGGGCGAGGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCA		
Db	1529 TTAAGAAGACACCTGGGCGAGGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCA		
QY	1021 ATATCATGTCCGCTTTGAAGCCCTTTCGGGCTTAATCTCAATGAA-TCAGAGCAGCA		
Db	1589 ATATCATGTCCGCTTTGAAGCCCTTTCGGGCTTAATCTCAATGAAAGTTCAGAGCAGCA		
QY	1080 ACTCATGGAATAATGATTCCTCTCTTTGAACGGGAGTTTGAGTACCAAGACCTTGATG		
Db	1649 ACTCTGGGAATAATGATTCCTCTCTTTGAACGGGAGTTTGAGTACCAAGACCTTGATG		
QY	1140 TGAAGAAGCATCGATGATGAAGATGTGGATGAAAAAGGATGATGATGGAATCAT 1199		
REFERENCE			
1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikawa, E., Oura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Hirakawa, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Hiyada, M., Houta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, P., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hiro, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Icho, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togaishi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs			
Nat. Genet. 36 (1), 40-45 (2004)			
14702039			
JOURNAL PUBMED			
REFERENCE			
2 Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Nabekura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y. and Oshima, A. NEDO human cDNA sequencing project			
Unpublished			
3 (bases 1 to 6045)			
Isogai, T. and Otsuki, T.			
Direct Submission			
Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)			
NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center			

etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES
Source
Location/Qualifiers
1..6045
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE4000128"
/tissue_type="placenta"
/clone_lib="PLACE4"
/note="cloning vector: pME18SFL3"
364..957
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14692.1"
/db_xref="GI:10435880"
/translation="MRVHCALIPSDMLHISNCRTABEMVLTLDDYLPHHEVQAVSSL
SQGKHGKQIDPITFYKIGIRLHYKFGITSDWYRIKQSIDSKRTAWRRKQWGQSL
VVKFSFRMPNSSSYCPSEPMWSTPPAPRELLOPQLOPQALHYALANVQVQIHQIG
DQOVQVPOGHLHIAKVPOGBOVQITQDSBSPSGTT"

ORIGIN
Query Match 41.6%; Score 792.8; DB 9; Length 6045;
Best Local Similarity 96.8%; Pred. No. 2.6e-185;
Matches 852; Conservative 0; Mismatches 22; Indels 6; Gaps 4;
QY 363 AGAGCCCGTGTCTGTACTAGTGAAGGAGTCTCAACAGCAGCAGCTGC-AGCGCT 421
DB 932 AAGCCCGTGTCTGTACTAGTGAAGGAGTCTCAACAGCAGCAGCTGCAGCGCT 991
QY 422 TCTACTCCTCGGCCACATCGCCTCAGACGTGGCGGGGTGGCGCTGGCGCTGTG 481
DB 992 TCTACTCCTCGGCCACATCGCCTCAGACGTGGCGGGGTGGCGCTGGCGCTGTG 1051
QY 482 CCCTCAACGACACTCCTCGAGCGCTACTCGACATGCTCTCGGCCGACCGCTGCAGGC 541
DB 1052 CCCTCAACGACACTCCTCGAGCGCTACTCGACATGCTCTCGGCCGACCGCTGCAGGC 1111
QY 542 TGAGCATTCTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGGTCCAGTATGCTTC 601
DB 1112 TCAGTACTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAGGTCCAGTATGCTTC 1171
QY 602 CTACATGGCAGCAGGTCTGAACCTCCATPACTCTTTGGCATTAACATCGACAAAGGATT 661
DB 1172 CTACATGGCAGCAGGTCCGAACCTCCATPACTCTTTGGCATTAACATCGACAAAGGATT 1231
QY 662 TGAACGGGCAGTAAGTTGCTCCGACCGTTTCAGACCTCTTAAGAGTCAACGAGA 721
DB 1232 TGAACGGGCAGTAAGTTGCTCCGACCGTTTCAGACCTCTTAAGAGTCAACGAGA 1291
QY 722 ACGTGACCTCCTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCTGTTCAAGGAGATCA 781
DB 1292 ATGTGA---CCTTGCTGAAGGAGTCCACGCAAGGAGTGAGCAGCTGTTCAAGGAGATCA 1348
QY 782 CAGCCTCTCTGCGCGTCTCCATCTCATCAAACTGAACAGGAGACCGA-CCCTTGCGCTG 840
DB 1349 CAGCCTCTCTGCGCATCTCCATCTCATCAAACTGAACAGGAGACCGA-CCCTTGCGCG 1408
QY 841 TCGTGTCCAGGAATGTCAGTCTGATGCCAATTCGAAAGAGGAGCGGAGAGAAAGA 900
DB 1409 TCGTGTCCAGGAATGTCAGTCTGATGCCAATTCGAAAGAGGAGCGGAGAGAAAGC 1468
QY 901 AAGTGACCAACATATCTCATTTTGTATGAGGAAGATGAGCAGAACTCTGGGACCTGT 960
DB 1469 AAGTGACCAACATATCTCATTTGTATGAGGAAGATGAGCAGAACTCTGGGACATGT 1528
QY 961 TTAAGAGACACCTGGGGCAGGAGAGCTCAGAGGACAACTCGACCGCTCTCTGTCTCA 1020
DB 1529 TTAAGAGACACCTGGGGCAGGAGAGCTCAGAGGACAACTCGACCACTCTCTGTCTCA 1588
QY 1021 ATATCATGTCCGCTTTGAAGAGCCCTTCGGGCTTAACCTCAATGGAA-TCAGAGGAGCA 1079
DB 1589 ATATCATGTCCGCTTTGAAGAGCCCTTCGGGCGAAACTCAATGGGAAGTCAGAGGAGCA 1648

QY 1080 ACTCATGAAATTTGATTCCCTGTCTTTGAACGGGAGTTTGGTACCAGAACTTGATG 1139
DB 1649 ACTCGTGAATTTGATTCCCTGTCTTTGAACGGGAGTTTGGTACCAGAACTTGATG 1708
QY 1140 TGAAGACATCGATGAAAGATGTCGATGAAAGAGATGACGTGTATGGAACATCAT 1199
DB 1709 TGAAGACATCGATGAAAGATGTCGATGAAAGAGATGACGTGTATGGAACATCAT 1768
QY 1200 CAGGAAGAGACACAGGGGCCACTCGGAGTCGCCCGAGAA 1239
DB 1769 CAGGACGGAAGACACAGGGGCCACTCAGAGTCGCCCGAGAA 1808

RESULT 13
BD146965
LOCUS
DEFINITION
Primer for synthesizing full-length cDNA and use thereof.
ACCESSION
BD146965
VERSION
BD146965.1 GI:27852723
KEYWORDS
JP 2002191363-A/1808.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 730)
Oka,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 1808 09-JUL-2002;
JOURNAL
HELIX RESEARCH INSTITUTE
COMMENT
OS Homo sapiens (human)
PN JP 2002191363-A/1808
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
10,
PC C12P21/02,C12Q1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof PH Key
Location/Qualifiers
FT source 1..730
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
Location/Qualifiers
1..730
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 35.3%; Score 672.2; DB 6; Length 730;
Best Local Similarity 97.9%; Pred. No. 1.9e-155;
Matches 696; Conservative 0; Mismatches 13; Indels 2; Gaps 2;
QY 77 GCGCGCGCGCGCGCGCANGCACCGGCCCGGGANAGCNCATGAGCGGATCNGAACN 136
DB 22 GCGCGCGCGCGCGCGCANGCACCGGCCCGGGAGAGGCACCATGAGCGGATCAAGAAC 81
QY 137 ATGACAAAAGACAAATTTCTGCTGGAGCGACTGCTGGATGACGTGAACAGTGCAGATCC 196
DB 82 ATGACAAAAGACAAATTTCTGCTGGAGCGACTGCTGGATGACGTGAACAGTGCAGATCC 141
QY 197 GCTTTNAGGGAGAAAGGAGATTGCTCGGATTCGACAGCAGGGTCACTGCTGTGTG 256
DB 142 GCTTTNAGGGAGAAAGGAGATTGCTCGGATTCGACAGCAGGGTCACTGCTGTGTG 201
QY 257 CCAGTTTGAAGCCCTTCCTGCAGCAGTGGCTTGAAGAGGAGTGCAGATTCGACAG 316
DB 202 CCAGTTTGAAGCCCTTCCTGCAGCAGTGGCTTGAAGAGGAGTGCAGATTCGACAG 261

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:57:41 ; Search time 47.1575 Seconds
(without alignments)
2390.553 Million cell updates/sec

Title: US-09-155-676B-7

Perfect score: 5052
Sequence: 1 MAVMEMACPGAGSAGVQK.....PDGSFAWSRWKVGQLENRP 947

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 557933 seqs, 119041758 residues

Total number of hits satisfying chosen parameters: 557933

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New:*

- 1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US11_NEW_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5052	100.0	947	5	US-09-155-676B-7
2	5052	100.0	947	5	US-09-155-676B-14
3	5040	99.8	947	8	US-60-659-397-1471
4	5040	99.8	947	8	US-60-659-397-1472
5	391.5	7.7	1673	8	US-60-643-717-14226
6	367	7.3	256	7	US-11-113-424-183
7	360	7.1	467	8	US-60-659-397-1644
8	360	7.1	475	6	US-10-940-774A-10456
9	359	7.1	467	5	US-09-155-676B-13
10	358	7.1	351	1	PCT-US04-30360-157
11	358	7.1	351	6	US-10-941-635-157
12	355.5	7.1	467	5	US-09-155-676B-12
13	355.5	7.0	672	5	US-09-155-676B-19
14	355.5	7.0	672	5	US-09-608-890A-2
15	355.5	7.0	1593	5	US-09-608-890A-4
16	353	7.0	619	5	US-09-608-890A-6
17	352	7.0	619	5	US-09-608-890A-8
18	349	6.9	1722	8	US-60-643-717-11386
19	346	6.8	1495	6	US-10-450-763-34736
20	343.5	6.8	614	8	US-60-643-717-691
21	341	6.7	1320	8	US-60-643-717-10799
22	338.5	6.7	611	8	US-60-643-717-689
23	337.5	6.7	314	1	PCT-US04-30360-120
24	337.5	6.7	314	6	US-10-941-635-120
25	325	6.4	1478	5	US-09-155-676B-18

ALIGNMENTS

RESULT 1

US-09-155-676B-7

; Sequence 7, Application US/09155676B

; GENERAL INFORMATION:

; APPLICANT: WALLACH, David

; APPLICANT: MALININ, Nikolai

; APPLICANT: BOLDIN, Mark

; APPLICANT: KOVALENKO, Andrei

; APPLICANT: METT, Igor

; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR

; FILE REFERENCE: WALLACH=21

; CURRENT APPLICATION NUMBER: US/09/155.676B

; PRIOR FILING DATE: 1999-01-04

; PRIOR APPLICATION NUMBER: PCT/IL97/00117

; PRIOR FILING DATE: 1997-04-01

; PRIOR APPLICATION NUMBER: IL 117800

; PRIOR FILING DATE: 1996-04-02

; PRIOR APPLICATION NUMBER: IL 119133

; PRIOR FILING DATE: 1996-08-26

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 7

; LENGTH: 947

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-155-676B-7

Query Match 100.0%; Score 5052; DB 5; Length 947;

Best Local Similarity 100.0%; Pred. No. 6.9e-231;

Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAVMEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Db	1	MAVMEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Qy	61	VITKGTAKGSEAGPAAISIIAQACENSQSFPTFSERIFIAGSKQYSQSESLDQIPNN 120
Db	61	VITKGTAKGSEAGPAAISIIAQACENSQSFPTFSERIFIAGSKQYSQSESLDQIPNN 120
Qy	121	VAHATEGMARVCKGGRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180
Db	121	VAHATEGMARVCKGGRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180
Qy	181	DESLPGAPYVNTPQFTKPLKEGLGQICFKQGEGLRPAIPRSELHKLISPLQCLNHW 240
Db	181	DESLPGAPYVNTPQFTKPLKEGLGQICFKQGEGLRPAIPRSELHKLISPLQCLNHW 240

Sequence 9386, Ap
Sequence 16140, A
Sequence 16764, A
Sequence 41304, A
Sequence 4, Appli
Sequence 11116, A
Sequence 8279, Ap
Sequence 10942, A
Sequence 451, App
Sequence 10, Appl
Sequence 11014, A
Sequence 25, Appl
Sequence 136408,
Sequence 508, App
Sequence 912, App
Sequence 13971, A
Sequence 60, Appl
Sequence 1857, Ap
Sequence 6680, Ap
Sequence 1594, Ap

26 325 6.4 1478 8 US-60-643-717-9386
27 325 6.4 1478 8 US-60-643-717-16140
28 325 6.4 1478 8 US-60-643-717-16764
29 324.5 6.4 1612 7 US-11-097-143-41304
30 322 6.4 1315 1 PCT-US03-27894-4
31 318.5 6.3 1955 8 US-60-643-717-11116
32 315.5 6.2 1423 8 US-60-643-717-8279
33 313 6.2 624 6 US-10-940-774A-10942
34 313 6.2 626 1 PCT-US03-10870-451
35 313 6.2 626 5 US-09-608-890A-10
36 307 6.1 1832 8 US-60-643-717-11014
37 305.5 6.0 525 6 US-10-491-545A-25
38 304.5 6.0 522 6 US-10-703-032-136408
39 304.5 6.0 965 7 US-11-090-997-508
40 304.5 6.0 1490 1 PCT-US03-10870-912
41 304.5 6.0 2139 7 US-11-097-143-13971
42 303 6.0 2523 6 US-10-130-978-60
43 301.5 6.0 968 1 PCT-US04-42360-1857
44 301.5 6.0 968 6 US-10-940-774A-6680
45 301.5 6.0 974 7 US-11-090-997-1594


```

QY 241 KLHPDQGGPLPLTHPPFYSRPLPHPPHPLPQWKPHPLESFLGKLACVDSOKPLPDPH 300
Db 241 KLHPDQGGPLPLTHPPFYSRPLPHPPHPLPQWKPHPLESFLGKLACVDSOKPLPDPH 300
QY 301 LSKLACVDSGPKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
Db 301 LSKLACVDSGPKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
QY 361 AARGSRREPSKPTEDNEGVLITLTKLPVDYEVREEVHATHQRLGRSGFGEVHRMEDK 420
Db 361 AARGSRREPSKPTEDNEGVLITLTKLPVDYEVREEVHATHQRLGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAELMACAGLTSPIVPLYGAVREGPWVNI FMELEGGSLGOL 480
Db 421 QTGFQCAVKVRLEVFRAELMACAGLTSPIVPLYGAVREGPWVNI FMELEGGSLGOL 480
QY 481 VKEQGCLPEDRALYYLGQALEGLEYLHRSRRIILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGCLPEDRALYYLGQALEGLEYLHRSRRIILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGKSLITGDYIPGTETHMAPEVNLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Db 541 LQPDGLGKSLITGDYIPGTETHMAPEVNLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
QY 601 RGPLCLKIASBPVPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAELGKGNRALQOQVG 660
Db 601 RGPLCLKIASBPVPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAELGKGNRALQOQVG 660
QY 661 LKSPMRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPPRAEETTGAPKLQPLPPE 720
Db 661 LKSPMRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPPRAEETTGAPKLQPLPPE 720
QY 721 PEPNKSPPPLTISKESGMEWPEPLSSLEPAPARNPSSPERKATVPQELOLELEFLN 780
Db 721 PEPNKSPPPLTISKESGMEWPEPLSSLEPAPARNPSSPERKATVPQELOLELEFLN 780
QY 781 SLSQPFSLBEEQIILSCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSSQAEARSS 840
Db 781 SLSQPFSLBEEQIILSCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSSQAEARSS 840
QY 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
Db 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
QY 901 LVTXGQPVRYDMEVPDSDGIDLQCTLAPDGSFAWSWRVKHGOLENRP 947
Db 901 LVTXGQPVRYDMEVPDSDGIDLQCTLAPDGSFAWSWRVKHGOLENRP 947

RESULT 2
US-09-155-676B-14
; GENERAL INFORMATION: US/09155676B
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAP), THEIR
; PREPARATION AND USE
; FILE REFERENCE: WALLACH-21
; CURRENT APPLICATION NUMBER: US/09/155,676B
; CURRENT FILING DATE: 1999-01-04
; PRIOR APPLICATION NUMBER: PCT/IL97/00117
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: IL 117800
; PRIOR FILING DATE: 1996-04-02
; PRIOR APPLICATION NUMBER: IL 119133
; PRIOR FILING DATE: 1996-08-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14
; LENGTH: 947
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-155-676B-14

Query Match      100.0%; Score 5052; DB 5; Length 947;
Best Local Similarity 100.0%; Pred. No. 6.9e-231;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVMEMACGAPGSAVGQOKELPKPKETPTPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACGAPGSAVGQOKELPKPKETPTPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITTKTAKGSGAGAAISITAAACENSQESPTFSERIFIAGSKQVSSQESLQIPNN 120
Db 61 VITTKTAKGSGAGAAISITAAACENSQESPTFSERIFIAGSKQVSSQESLQIPNN 120
QY 121 VAHATEGKMARVCWKGRSKARKKKSSKSLAHAGVALAKPLPRTEQESCTIPVQE 180
Db 121 VAHATEGKMARVCWKGRSKARKKKSSKSLAHAGVALAKPLPRTEQESCTIPVQE 180
QY 181 DESPLGAPYVRNTPQFTKPLKEPGLGOLCFKOLGGLRALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPYVRNTPQFTKPLKEPGLGOLCFKOLGGLRALPRSELHKLISPLQCLNHVW 240
QY 241 KLHHPDQGGPLPLTHPPFPYSRPLPHPPHPLPQWKPHPLESFLGKLACVDSOKPLPDPH 300
Db 241 KLHHPDQGGPLPLTHPPFPYSRPLPHPPHPLPQWKPHPLESFLGKLACVDSOKPLPDPH 300
QY 301 LSKLACVDSGPKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
Db 301 LSKLACVDSGPKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
QY 361 AARGSRREPSKPTEDNEGVLITLTKLPVDYEVREEVHATHQRLGRSGFGEVHRMEDK 420
Db 361 AARGSRREPSKPTEDNEGVLITLTKLPVDYEVREEVHATHQRLGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAELMACAGLTSPIVPLYGAVREGPWVNI FMELEGGSLGOL 480
Db 421 QTGFQCAVKVRLEVFRAELMACAGLTSPIVPLYGAVREGPWVNI FMELEGGSLGOL 480
QY 481 VKEQGCLPEDRALYYLGQALEGLEYLHRSRRIILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGCLPEDRALYYLGQALEGLEYLHRSRRIILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGKSLITGDYIPGTETHMAPEVNLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Db 541 LQPDGLGKSLITGDYIPGTETHMAPEVNLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
QY 601 RGPLCLKIASBPVPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAELGKGNRALQOQVG 660
Db 601 RGPLCLKIASBPVPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAELGKGNRALQOQVG 660
QY 661 LKSPMRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPPRAEETTGAPKLQPLPPE 720
Db 661 LKSPMRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPPRAEETTGAPKLQPLPPE 720
QY 721 PEPNKSPPPLTISKESGMEWPEPLSSLEPAPARNPSSPERKATVPQELOLELEFLN 780
Db 721 PEPNKSPPPLTISKESGMEWPEPLSSLEPAPARNPSSPERKATVPQELOLELEFLN 780
QY 781 SLSQPFSLBEEQIILSCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSSQAEARSS 840
Db 781 SLSQPFSLBEEQIILSCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSSQAEARSS 840
QY 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
Db 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
QY 901 LVTXGQPVRYDMEVPDSDGIDLQCTLAPDGSFAWSWRVKHGOLENRP 947
Db 901 LVTXGQPVRYDMEVPDSDGIDLQCTLAPDGSFAWSWRVKHGOLENRP 947
```



```
RESULT 3
US-60-659-397-1471
; Sequence 1471, Application US/60659397
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/659,397
; CURRENT FILING DATE: 2005-03-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1471
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-659-397-1471

Query Match      99.8%; Score 5040; DB 8; Length 947;
Best Local Similarity 99.8%; Pred. No. 2.6e-230;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAVMEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
DB 1 MAVMEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKGSEAGPAAISIIAQACENSQSFPTFSERIFIAGSKQYSQESLDQIPNN 120
DB 61 VITKGTAKGSEAGPAAISIIAQACENSQSFPTFSERIFIAGSKQYSQESLDQIPNN 120
QY 121 VAHATEGMARVCWKGRSKARKKRSKSLAHAGVALAKPLPRTEOESCTIPVOE 180
DB 121 VAHATEGMARVCWKGRSKARKKRSKSLAHAGVALAKPLPRTEOESCTIPVOE 180
QY 181 DESPLGAPYVRNTPQFTKPLKEPGLGQCFKQGLGRLPALPRSELHKLISPLQCLNHV 240
DB 181 DESPLGAPYVRNTPQFTKPLKEPGLGQCFKQGLGRLPALPRSELHKLISPLQCLNHV 240
QY 241 KLHPDQGGPLPLTPHPFPYSRLPHPPFHPQPKHPLESFLGKLACVDSQKPLPDHP 300
DB 241 KLHPDQGGPLPLTPHPFPYSRLPHPPFHPQPKHPLESFLGKLACVDSQKPLPDHP 300
QY 301 LSKLACVDSQKPLPDHPLESCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
DB 301 LSKLACVDSQKPLPDHPLESCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
QY 361 AARGSRSEPSPKTEDNEGVLITTEKLPVDYVEEVEVHWATHQLRGRSGFGEVHRMEDK 420
DB 361 AARGSRSEPSPKTEDNEGVLITTEKLPVDYVEEVEVHWATHQLRGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVLVYAVREGPWNIIPWELLEGSLGOL 480
DB 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVLVYAVREGPWNIIPWELLEGSLGOL 480
QY 481 VKEQGLCEDRALYVLGQALSGLEVILHGRRIILHGDVKAADNVLLSDGSHAALCDFGHAVC 540
DB 481 VKEQGLCEDRALYVLGQALSGLEVILHGRRIILHGDVKAADNVLLSDGSHAALCDFGHAVC 540
QY 541 LQPDGLGKSLITGDYIPGTETHMAPEVVLGRSCDAKVDWVSSCCMMLHMLNGCHPWTQFF 600
DB 541 LQPDGLGKSLITGDYIPGTETHMAPEVVLGRSCDAKVDWVSSCCMMLHMLNGCHPWTQFF 600
QY 601 RGPLCLKIASPPPPVREIPPSCAPITAQAIQEGLRKEPIHRVSAEALGGKVNRLAQVGG 660
DB 601 RGPLCLKIASPPPPVREIPPSCAPITAQAIQEGLRKEPIHRVSAEALGGKVNRLAQVGG 660
QY 661 LKSPWRGEYKEPRHPHPNOANYHOTLHAQPRELSPRAPCPRAEETTGRAPKLPPLPPE 720
DB 661 LKSPWRGEYKEPRHPHPNOANYHOTLHAQPRELSPRAPCPRAEETTGRAPKLPPLPPE 720

RESULT 4
US-60-659-397-1472
; Sequence 1472, Application US/60659397
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/659,397
; CURRENT FILING DATE: 2005-03-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1472
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-659-397-1472

Query Match      99.8%; Score 5040; DB 8; Length 947;
Best Local Similarity 99.8%; Pred. No. 2.6e-230;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAVMEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
DB 1 MAVMEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKGSEAGPAAISIIAQACENSQSFPTFSERIFIAGSKQYSQESLDQIPNN 120
DB 61 VITKGTAKGSEAGPAAISIIAQACENSQSFPTFSERIFIAGSKQYSQESLDQIPNN 120
QY 121 VAHATEGMARVCWKGRSKARKKRSKSLAHAGVALAKPLPRTEOESCTIPVOE 180
DB 121 VAHATEGMARVCWKGRSKARKKRSKSLAHAGVALAKPLPRTEOESCTIPVOE 180
QY 181 DESPLGAPYVRNTPQFTKPLKEPGLGQCFKQGLGRLPALPRSELHKLISPLQCLNHV 240
DB 181 DESPLGAPYVRNTPQFTKPLKEPGLGQCFKQGLGRLPALPRSELHKLISPLQCLNHV 240
QY 241 KLHPDQGGPLPLTPHPFPYSRLPHPPFHPQPKHPLESFLGKLACVDSQKPLPDHP 300
DB 241 KLHPDQGGPLPLTPHPFPYSRLPHPPFHPQPKHPLESFLGKLACVDSQKPLPDHP 300
QY 301 LSKLACVDSQKPLPDHPLESCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
DB 301 LSKLACVDSQKPLPDHPLESCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
QY 361 AARGSRSEPSPKTEDNEGVLITTEKLPVDYVEEVEVHWATHQLRGRSGFGEVHRMEDK 420
DB 361 AARGSRSEPSPKTEDNEGVLITTEKLPVDYVEEVEVHWATHQLRGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVLVYAVREGPWNIIPWELLEGSLGOL 480
DB 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVLVYAVREGPWNIIPWELLEGSLGOL 480
QY 481 VKEQGLCEDRALYVLGQALSGLEVILHGRRIILHGDVKAADNVLLSDGSHAALCDFGHAVC 540
DB 481 VKEQGLCEDRALYVLGQALSGLEVILHGRRIILHGDVKAADNVLLSDGSHAALCDFGHAVC 540
QY 541 LQPDGLGKSLITGDYIPGTETHMAPEVVLGRSCDAKVDWVSSCCMMLHMLNGCHPWTQFF 600
DB 541 LQPDGLGKSLITGDYIPGTETHMAPEVVLGRSCDAKVDWVSSCCMMLHMLNGCHPWTQFF 600
QY 601 RGPLCLKIASPPPPVREIPPSCAPITAQAIQEGLRKEPIHRVSAEALGGKVNRLAQVGG 660
DB 601 RGPLCLKIASPPPPVREIPPSCAPITAQAIQEGLRKEPIHRVSAEALGGKVNRLAQVGG 660
QY 661 LKSPWRGEYKEPRHPHPNOANYHOTLHAQPRELSPRAPCPRAEETTGRAPKLPPLPPE 720
DB 661 LKSPWRGEYKEPRHPHPNOANYHOTLHAQPRELSPRAPCPRAEETTGRAPKLPPLPPE 720
```


QY	558	CTETHMAPEVLGRSCDAKVDVWSSCCMMLHMLNCHCHPWTOFF---RGPLCLKIA-SEPP	613
Db	289	GTEIYMSPEVILCRGHSTKADIYSLGATLIHMOTGTPPWVKYPRSAVPSYLIHKOAP	348
QY	614	PVREIPPSCAPLTAQAIQEGLRKEPIHRVSAALGGKVNRAALQQVGGGLKSPMRGEYKEPR	673
Db	349	PLEDIADDCSPGMRELIEASLERNPNRPRADL-----LK-----	384
QY	674	HPPPNQANYHTLHAQPRELSPRAPGPRPAEETTGRAKLPQPLPPEPPEPNKSPPLTSL	733
Db	385	-----HEALN-PPREDQPRC---OSLSALLERKRLLSRKELELPENIADSSCTGS	431
QY	734	KEESGM 739	
Db	432	TEESEM 437	
RESULT 8			
US-10-940-774A-10456			
; Sequence 10456, Application US/10940774A			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
; FILE REFERENCE: CL001307			
; CURRENT APPLICATION NUMBER: US/10/940,774A			
; CURRENT FILING DATE: 2004-09-15			
; PRIOR APPLICATION NUMBER: 60/241,755			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 60/237,768			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/231,498			
; PRIOR FILING DATE: 2000-09-08			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 10456			
; LENGTH: 475			
; TYPE: PRT			
; ORGANISM: Human			
US-10-940-774A-10456			
Query Match 7.1%; Score 360; DB 6; Length 475;			
Best Local Similarity 27.2%; Pred. No. 9.2e-10;			
Matches 116; Conservative 69; Mismatches 155; Indels 86; Gaps 14;			
QY	332	SVEEYLVHALQGSVSSQASHLTSIAKTWAARGSRSPSKTNEGVLLTEKLKPDY	391
Db	88	TVEDLL--AFANHISNTAKHY-----GQRPQE-----SGILLNMVITPQNG	127
QY	392	EYREE-----VHWATHQLRLG-----RGSFGVHMEKDQTCFOCAVKVRLVFAEEL	441
Db	128	RYQISDVLLIPWKLTYRNIGSDFTPRGAFGVYLAQDIKTKRWACKLIIPVDQPKPSDV	187
QY	442	MACAGLTSPIRVLYGAVREGPWVNIIFMELLEGSGSLQGVKEQGLCPEDRALYVLGQALE	501
Db	188	EQACFRHENIAELYGAVLWGETVHLFWEAGEGGSVLEKLESCGPMREFEIIWTKVHLK	247
QY	502	GLEYLHSRILHGDVKADVNLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDIYP	557
Db	248	GLDFLHSKVIHDIKPSNIVFMS--TKAVLVDFGLSVQMTED-----VYFPKDLR	296
QY	558	GTETHMAPEVLGRSCDAKVDVWSSCCMMLHMLNCHCHPWTOFF---RGPLCLKIA-SEPP	613
Db	297	GTEIYMSPEVILCRGHSTKADIYSLGATLIHMOTGTPPWVKYPRSAVPSYLIHKOAP	356
QY	614	PVREIPPSCAPLTAQAIQEGLRKEPIHRVSAALGGKVNRAALQQVGGGLKSPMRGEYKEPR	673
Db	357	PLEDIADDCSPGMRELIEASLERNPNRPRADL-----LK-----	392
QY	674	HPPPNQANYHTLHAQPRELSPRAPGPRPAEETTGRAKLPQPLPPEPPEPNKSPPLTSL	733
Db	393	-----HEALN-PPREDQPRC---OSLSALLERKRLLSRKELELPENIADSSCTGS	439

Qy 734 KEESGM 739
Db 440 TESEEM 445

RESULT 9

US-09-155-676B-13

; Sequence 13, Application US/09155676B

; GENERAL INFORMATION:

; APPLICANT: WALLACH, David

; APPLICANT: MALININ, Nikolai

; APPLICANT: BOLDIN, Mark

; APPLICANT: KOVALENKO, Andrei

; APPLICANT: METT, Igor

; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR

; TITLE OF INVENTION: PREPARATION AND USE

; FILE REFERENCE: WALLACH-21

; CURRENT APPLICATION NUMBER: US/09/155,676B

; CURRENT FILING DATE: 1999-01-04

; PRIOR APPLICATION NUMBER: PCT/IL97/00117

; PRIOR FILING DATE: 1997-04-01

; PRIOR APPLICATION NUMBER: IL 117800

; PRIOR FILING DATE: 1996-04-02

; PRIOR APPLICATION NUMBER: IL 119133

; PRIOR FILING DATE: 1996-08-26

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 13

; LENGTH: 467

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-155-676B-13

Query Match

Best Local Similarity 7.1%; Score 359; DB 5; Length 467;

Matches 116; Conservative 69; Mismatches 154; Indels 88; Gaps 14;

Qy 332 SVEYLVHALQGSVSSQAHSLTSLAKTWAAGRSRSPKTDNEGVLLEKLPVDY 391
Db 80 TVEDLL--AFANHSNTAKHFY-----GQRPQE-----SGILLNNVITPQNG 119

Qy 392 EYREE-----VHWATHQLRLG-----RGSFGEVHRMEDKQTGFQCAVKVRLEVFRABEL 441
Db 120 RYQIDSDVLLIPMKLTYRNIGSDFIPIRGAFGVYLAQDIKTKRMACKLIPVDQFKPSDV 179

Qy 442 MACAGLTSPIRVLYGAVREGPWNIIFMELLEGSLGQLVKEQCLPEDRALYVLGQALE 501
Db 180 ETQACFRHENIAELYGAVLWGETVHLFMEAGEGGVLEKLESCGPMREFEIIWTKHVLIK 239

Qy 502 GLEYLHRRILHGDVKADNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLITGDYIP---- 557
Db 240 GLDFLHKKVVIHHDIKPSNIVFWS--TKAVLVDFGLSV-----QMTEDVYFPKDL 287

Qy 558 -GTETHMAPEVVLGRSCDAKVDVWSSCCMLHMLNGCHPWTOFF---RGPLCLKTA-SEP 612
Db 288 RTEIYMSPEVILCRGHSTKADIYSLGATLIHMOTGTPPWVKRYPRSAYPEYLYIHKQA 347

Qy 613 PVPREIPSPCAPLTAQAIQEGRLKSPPIHRVSAALGGKVNRAALQVGGKLSWRCGEYKEP 672
Db 348 PPLEDIADDCSPGMRELIEASLERPNHRPRAADL-----LK----- 384

Qy 673 RHPPNQANYHTLHAQPRELSPAGRPRAEETGTRAPKLQPPLPPEPPPNKSPPLTL 732
Db 385 -----HEALN-PPREDQPRC---QSLDSALLERKRLLSRKELELPENIADSSCTG 430

Qy 733 SKEESGM 739

Db 431 STEESEM 437

RESULT 10

PCT-US04-30360-157

; Sequence 157, Application PC/TUS0430360

; GENERAL INFORMATION:

; APPLICANT: PLEXIKON, INC.

; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT

; FILE REFERENCE: 039363-1703

; CURRENT APPLICATION NUMBER: PCT/US04/30360

; CURRENT FILING DATE: 2004-09-15

; PRIOR APPLICATION NUMBER: 60/503,277

; PRIOR FILING DATE: 2003-09-15

; NUMBER OF SEQ ID NOS: 167

; SOFTWARE: PatentIn Ver. 3.2

; SEQ ID NO 157

; LENGTH: 351

; TYPE: PRT

; ORGANISM: Unknown Organism

; FEATURE:

; OTHER INFORMATION: Description of Unknown Organism: Mammalian

; OTHER INFORMATION: protein sequence

PCT-US04-30360-157

Query Match

Best Local Similarity 7.1%; Score 358; DB 1; Length 351;

Matches 113; Conservative 68; Mismatches 151; Indels 76; Gaps 13;

Qy 350 AHSLSLAKTWAAGRSRSPKTDNEGVLLEKLPVDYVEYREE-----VHWATHQL 404
Db 1 ANHSNTAKHP--YQRPQE-----SGILLNNVITPQNGRYQIDSDVLLIPMKLTYR 50

Qy 405 RLG-----RGSFGEVHRMEDKQTGFQCAVKVRLEVFRAEELMACAGLTSPIRVLYGAV 459
Db 51 NIGSDFIPIRGAFGVYLAQDIKTKRMACKLIPVDQFKPSDVEIQACFRHENIAELYGAV 110

Qy 460 REGPWNIIFMELLEGSLGQLVKEQCLPEDRALYVLGQALEGLYHLSRLLHGDVKAD 519
Db 111 LWGETVHLFMEAGEGGVLEKLESCGPMREFEIIWTKHVLLKGLDFLHKKVVIHHDIKPS 170

Qy 520 NVLLSSDGSAAALCDFGHAVCLQPDGLGKSLITGDYIP---GTETHMAPEVVLGRSCDA 575
Db 171 NIVFWS--TKAVLVDFGLSVQMTED-----VYFPKDLRGTEIYMSPEVILCRGHST 219

Qy 576 KVDVWSSCCMLHMLNGCHPWTOFF---RGPLCLKIA-SEPPVPREIPSPCAPLTAQAIQ 631
Db 220 KADIYSLGATLIHMOTGTPPWVKRYPRSAYPEYLYIHKQAPPLIEDIADDCSPGMRELIE 279

Qy 632 EGLRKEPIHRVSAALGGKVNRAALQVGGKLSWRCGEYKEPPEPPPNQANYHOTLHAQPR 691
Db 280 ASLERPNHRPRAADL-----LK-----HEALN-PPR 305

Qy 692 ELSRAPGRPAEETGTRAPKLQPPLPPEPPPNKSPPLTLKSKEESGM 739
Db 306 EDQPRCTS---LDSALLERKRLLSRKELELPENIADSSCTGSTEESEM 350

RESULT 11

US-10-941-635-157

; Sequence 157, Application US/10941635

; GENERAL INFORMATION:

; APPLICANT: ARTIS, DEAN R.

; APPLICANT: BREWER, RYAN E.

; APPLICANT: GILLETTE, SAMUEL J.

; APPLICANT: HURT, CLARENCE R.

; APPLICANT: IBRAHIM, PRABHA L.

; APPLICANT: ZUCKERMAN, REBECCA L.

; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT

; FILE REFERENCE: 039363-1702

; CURRENT APPLICATION NUMBER: US/10/941,635

; CURRENT FILING DATE: 2004-09-15

; PRIOR APPLICATION NUMBER: 60/503,277

; PRIOR FILING DATE: 2003-09-15

; NUMBER OF SEQ ID NOS: 167

; SOFTWARE: PatentIn Ver. 3.2

; SEQ ID NO 157

; LENGTH: 351

; TYPE: PRT

Db 249 SRCDDSGGGNSGNNAVIPSDETFTVPEVKCRDLVNTLNSSIEDLLEASMPSSDITVT 308
QY 348 -----SOAHSLSLAKTWAARGSR--EPSP 372
Db 309 FKSEVAVLSPEKAENDTYKDDVNNHOKCKEKEAEEAALAMAMASQDALPIVPOL 368
QY 373 KTDNEGVLLEKLP-----VDYEREVHATHQLRGRSGFGEVHRMEDKQTGF 424
Db 369 QVENGEDIQDITPETLPHTKAKOPYREDAEWLKGQ-QTGLGAFSSCYQAQDVGTGT 427
QY 425 QCAVKV-----RLEVFA--BELMACAGLTSRIVPLYGAVREGPWNIFMELLE 473
Db 428 LMAVKQVTVYRNTSSQEVEVEALREERMMGHLPNIRMLGATCEKSNYNLFIEWMA 487
QY 474 GSGLGOLVKEQCLPEDRALYLGQALEGLYLSRILHGDVKNVLLSSDGSAAALC 533
Db 488 GGSVAHLLSKYGAFKESVINYTEQLRGLSYLHENQIIRHDVKGANLLIDSTGQRLRIA 547
QY 534 DFGHAVCLQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGC 593
Db 548 DFGAAARLASKGTGAGEFQGLL-GTIAFMAPEVLGQYGRSCDVWSVGCALIEMACAK 606
QY 594 HPWTQFRG---PLCLKIASE-----PPVREIPSPCAPLTAQ--AIQEGLRKEP 638
Db 607 PPNNAEKHNHLALIFKIASATTAPSPHLSFGLRDVAVRCLELQPDORPPSRELLKHP 666
QY 639 IHRVS 643
Db 667 VERTT 671

RESULT 14
US-09-608-890A-2
; Sequence 2, Application US/09608890A
; GENERAL INFORMATION:
; APPLICANT: Johnson, Gary L.
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To
; TITLE OF INVENTION: External Signals
; FILE REFERENCE: CPI-004DVC3CN
; CURRENT APPLICATION NUMBER: US/09/608,890A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 08/440,421
; PRIOR FILING DATE: 1993-05-15
; PRIOR APPLICATION NUMBER: 08/323,460
; PRIOR FILING DATE: 1994-10-14
; PRIOR APPLICATION NUMBER: 08/410,602
; PRIOR FILING DATE: 1993-04-15
; PRIOR APPLICATION NUMBER: 08/410,602
; PRIOR FILING DATE: 1995-03-24
; PRIOR APPLICATION NUMBER: 08/472,934
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-608-890A-2

Query Match 7.0%; Score 355.5; DB 5; Length 672;
Best Local Similarity 23.6%; Pred. No. 2.2e-09;
Matches 157; Conservative 81; Mismatches 278; Indels 149; Gaps 24;
QY 77 AISIIAQAECENSQBFSTFSERIFAGSKQYSQSESLDQIPNNVAHATEGMARVCWK 136
Db 58 SLQAVAPTSCLENSLSLEHTVHREKTKGLSATRLSASSEDISDLAGVSVGLPSSTTEQ 117
QY 137 KR---RSKARKRKXKSKSLAHAGVAKLPLRTPPEQSCITIPVOEDESPLGAPVNT 193
Db 118 PKPAVQTKGRPHSQCLNSPLSHA--QLMFPAPSAPCSSAPSV---DIS-----KHR 165
QY 194 PQFTKPLKEPGLQGLCFQKLBGLRPLPRSELHLKILSPLO--CLNHVWKLHHPDGGPL 251

Db 166 POAFVPCIPS-----ASPQO-RKFSLQFQNCSEH-----RDSOL 202
QY 252 PLPTHTPPYSLRPLPFPHPLOPKWPHLESFLGKLACVDQSKPLPDPHLSKLA----- 305
Db 203 ---SVFTQSRPPSPSNIRPKPSRPV-----GSTSKLGDATKSMTLDLGS 248
QY 306 --CVDS-----PKPLPGPH-----LEPSC--LSRGAHEKFSVEYLVLHALQGSVSS-- 347
Db 249 SRCDDSGGGNSGNNAVIPSDETFTVPEVKCRDLVNTLNSSIEDLLEASMPSSDITVT 308
QY 348 -----SOAHSLSLAKTWAARGSR--EPSP 372
Db 309 FKSEVAVLSPEKAENDTYKDDVNNHOKCKEKEAEEAALAMAMASQDALPIVPOL 368
QY 373 KTDNEGVLLEKLP-----VDYEREVHATHQLRGRSGFGEVHRMEDKQTGF 424
Db 369 QVENGEDIQDITPETLPHTKAKOPYREDAEWLKGQ-QTGLGAFSSCYQAQDVGTGT 427
QY 425 QCAVKV-----RLEVFA--EELMACAGLTSRIVPLYGAVREGPWNIFMELLE 473
Db 428 LMAVKQVTVYRNTSSQEVEVEALREERMMGHLPNIRMLGATCEKSNYNLFIEWMA 487
QY 474 GSGLGOLVKEQCLPEDRALYLGQALEGLYLSRILHGDVKNVLLSSDGSAAALC 533
Db 488 GGSVAHLLSKYGAFKESVINYTEQLRGLSYLHENQIIRHDVKGANLLIDSTGQRLRIA 547
QY 534 DFGHAVCLQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGC 593
Db 548 DFGAAARLASKGTGAGEFQGLL-GTIAFMAPEVLGQYGRSCDVWSVGCALIEMACAK 606
QY 594 HPWTQFRG---PLCLKIASE-----PPVREIPSPCAPLTAQ--AIQEGLRKEP 638
Db 607 PPNNAEKHNHLALIFKIASATTAPSPHLSFGLRDVAVRCLELQPDORPPSRELLKHP 666
QY 639 IHRVS 643
Db 667 VERTT 671

RESULT 15
US-09-608-890A-4
; Sequence 4, Application US/09608890A
; GENERAL INFORMATION:
; APPLICANT: Johnson, Gary L.
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To
; TITLE OF INVENTION: External Signals
; FILE REFERENCE: CPI-004DVC3CN
; CURRENT APPLICATION NUMBER: US/09/608,890A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 08/440,421
; PRIOR FILING DATE: 1995-05-15
; PRIOR APPLICATION NUMBER: 08/323,460
; PRIOR FILING DATE: 1994-10-14
; PRIOR APPLICATION NUMBER: 08/049,254
; PRIOR FILING DATE: 1993-04-15
; PRIOR APPLICATION NUMBER: 08/410,602
; PRIOR FILING DATE: 1995-03-24
; PRIOR APPLICATION NUMBER: 08/472,934
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1593
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-608-890A-4

Query Match 7.0%; Score 355.5; DB 5; Length 1593;
Best Local Similarity 23.6%; Pred. No. 5.5e-09;
Matches 157; Conservative 81; Mismatches 278; Indels 149; Gaps 24;
QY 77 AISIIAQAECENSQBFSTFSERIFAGSKQYSQSESLDQIPNNVAHATEGMARVCWK 136

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: June 7, 2005, 11:52:21 ; Search time 422.493 Seconds
(without alignments)
2618.037 Million cell updates/sec

Title: US-09-155-676B-7
Perfect score: 5052
Sequence: 1 MAYMEMACPGAPGSAVGQOK.....PDGSAFMSWRVHGQLENRP 947

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:
1: /cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US105_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
33: /cgn2_6/ptodata/1/paa/US107_COMB.pep.*
34: /cgn2_6/ptodata/1/paa/US108_COMB.pep.*
35: /cgn2_6/ptodata/1/paa/US109_COMB.pep.*
36: /cgn2_6/ptodata/1/paa/US110_COMB.pep.*
37: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5052	100.0	947	15	US-09-155-676-7
2	5052	100.0	947	15	US-09-155-676-14
3	5052	100.0	947	15	US-09-155-676A-7
4	5052	100.0	947	15	US-09-155-676A-14
5	5052	100.0	947	22	US-09-791-537-1606
6	5052	100.0	947	24	US-09-949-003C-2282
7	5052	100.0	947	25	US-09-981-397A-18
8	5052	100.0	947	29	US-10-394-322A-44
9	5052	100.0	947	37	US-60-366-892-44
10	5052	100.0	947	37	US-60-474-850-493
11	5040	99.8	947	26	US-10-087-192-888
12	5040	99.8	947	27	US-10-170-205B-8384
13	5040	99.8	947	27	US-10-170-205B-9184
14	5040	99.8	947	37	US-60-443-566-2590
15	5040	99.8	947	37	US-60-443-566-2591
16	5040	99.8	947	37	US-60-453-050-13077
17	5040	99.8	947	37	US-60-453-050-13078
18	5040	99.8	947	37	US-60-453-135-13077
19	5040	99.8	947	37	US-60-453-135-13078
20	5040	99.8	947	37	US-60-455-444-7013
21	5040	99.8	947	37	US-60-455-444-7014
22	5040	99.8	947	37	US-60-465-241-7013
23	5040	99.8	947	37	US-60-465-241-7014
24	5040	99.8	947	37	US-60-466-412-13077
25	5040	99.8	947	37	US-60-466-412-13078
26	5040	99.8	947	37	US-60-485-450-1471
27	5040	99.8	947	37	US-60-485-450-1472
28	5040	99.8	947	37	US-60-487-610-2325
29	5040	99.8	947	37	US-60-487-610-2326
30	5040	99.8	947	37	US-60-563-440-1471
31	5040	99.8	947	37	US-60-563-440-1472
32	5040	99.8	947	37	US-60-582-609-2325
33	5040	99.8	947	37	US-60-582-609-2326
34	5040	99.8	952	24	US-09-949-003C-3136
35	4946	97.9	939	37	US-60-248-798-348
36	4196.5	83.1	942	22	US-09-791-537-137083
37	3439.5	68.1	697	1	PCT-US03-26780-1811
38	2481.5	49.1	548	26	US-10-087-192-885
39	2301	45.5	477	33	US-10-760-320A-4878
40	2301	45.5	477	33	US-10-760-620A-4878
41	1572.5	31.1	368	22	US-09-758-457-758
42	1572.5	31.1	368	28	US-10-219-925-758
43	1028	20.3	188	37	US-60-195-052-734
44	992	19.6	184	37	US-60-196-711-1924
45	586	11.6	113	37	US-60-213-847-1073

ALIGNMENTS

RESULT 1
US-09-155-676-7
; Sequence 7, Application US/09155676
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
; FACTOR (TRAF), THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/155,676
FILING DATE: 04-JAN-1999
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IL97/00117
FILING DATE: 01-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 117800
FILING DATE: 02-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 119133
FILING DATE: 26-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH=21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 947 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-155-676-7

Query Match 100.0%; Score 5052; DB 15; Length 947;
Best Local Similarity 100.0%; Pred. No. 5.5e-306;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVMEMACPGAGSAGVQKQKELPKPKETPLGKQSSVYKLBKAVKSPVFCQKWLND 60
DB 1 MAVMEMACPGAGSAGVQKQKELPKPKETPLGKQSSVYKLBKAVKSPVFCQKWLND 60
QY 61 VITKGTAKGSEAGPAISIIAQACENSQEFSTPFIAGSKQYSQESLDOIINN 120
DB 61 VITKGTAKGSEAGPAISIIAQACENSQEFSTPFIAGSKQYSQESLDOIINN 120
QY 121 VAHATEGMAVCWKGRKRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180
DB 121 VAHATEGMAVCWKGRKRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180
QY 181 DESPLGAPYVRNTQFTKPLKEPGLGQLCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
DB 181 DESPLGAPYVRNTQFTKPLKEPGLGQLCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
QY 241 KLHHPDGGPLPLPHPPYRSLPPLPQPKPPLPQPKPPLPQPKPPLPQPKPPLPQPKP 300
DB 241 KLHHPDGGPLPLPHPPYRSLPPLPQPKPPLPQPKPPLPQPKPPLPQPKPPLPQPKP 300
QY 301 LSKLACVDSPKPLPGPHLEPSCLSRGAHEKTSVEYLVHVLQGVSSQASHLSLAKTW 360
DB 301 LSKLACVDSPKPLPGPHLEPSCLSRGAHEKTSVEYLVHVLQGVSSQASHLSLAKTW 360
QY 361 AARGSRSPSKTNEDEGVLITKLPVDYVEEVHWAHOLRLGRSGFGEVHRMEDK 420
DB 361 AARGSRSPSKTNEDEGVLITKLPVDYVEEVHWAHOLRLGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVKVLEVFRAELMACAGLTSPIVPLYGAVREGPWNIWFMLLEGSLGQL 480
DB 421 QTGFQCAVKVKVLEVFRAELMACAGLTSPIVPLYGAVREGPWNIWFMLLEGSLGQL 480
QY 481 VKEQGLPEDRALYYIGQALEGLYHSRRLHGVKADNVLLSSDGGSHAALCDFGHAVC 540
DB 481 VKEQGLPEDRALYYIGQALEGLYHSRRLHGVKADNVLLSSDGGSHAALCDFGHAVC 540

QY 541 LQPDGLKSLLTGDIYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNCHPWTQFF 600
DB 541 LQPDGLKSLLTGDIYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNCHPWTQFF 600
QY 601 RGPLCLCIASEPPVREIIPPSCAPLTAQAIQEGLRKEPIHRYVSAALGKGVNRALQOVGG 660
DB 601 RGPLCLCIASEPPVREIIPPSCAPLTAQAIQEGLRKEPIHRYVSAALGKGVNRALQOVGG 660
QY 661 LKSPWRGEYKEPRHPPPNOANYHTQTHAOPRELSPRAGPRPAEFTTGRAPKLOPPLPE 720
DB 661 LKSPWRGEYKEPRHPPPNOANYHTQTHAOPRELSPRAGPRPAEFTTGRAPKLOPPLPE 720
QY 721 PPEPKSPPLTLKSESGMWEPLPLSLSPAPARNPSSPERKATVPQEQLELFLN 780
DB 721 PPEPKSPPLTLKSESGMWEPLPLSLSPAPARNPSSPERKATVPQEQLELFLN 780
QY 781 SLQPFSLERQEQILSCLSDISLSDSEKKNPSKASQSRDTLSSGVHSSQAEARS 840
DB 781 SLQPFSLERQEQILSCLSDISLSDSEKKNPSKASQSRDTLSSGVHSSQAEARS 840
QY 841 SNNVTLARGPTPTPSYFNGVKVQIQLNGEHLHIREFHVKVGDIAIGISSQIPAAAFS 900
DB 841 SNNVTLARGPTPTPSYFNGVKVQIQLNGEHLHIREFHVKVGDIAIGISSQIPAAAFS 900
QY 901 LVTKDGPVRYDMEVPDSDGLDQCTLAPDGSFAMSWRVKHGQLENRP 947
DB 901 LVTKDGPVRYDMEVPDSDGLDQCTLAPDGSFAMSWRVKHGQLENRP 947

RESULT 2
US-09-155-676-14
Sequence 14, Application US/09155676
GENERAL INFORMATION:
APPLICANT: WALLACH, David
APPLICANT: MALININ, Nikolai
APPLICANT: BOLDIN, Mark
APPLICANT: KOVALENKO, Andrei
APPLICANT: METT, Igor
TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
STREET: 624 Ninth Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/155,676
FILING DATE: 04-JAN-1999
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IL97/00117
FILING DATE: 01-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 117800
FILING DATE: 02-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 119133
FILING DATE: 26-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH=21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 947 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-155-676-14

Query Match 100.0%; Score 5052; DB 15; Length 947;
Best Local Similarity 100.0%; Pred. No. 5.5e-306;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVMEMCPGAGSAGVQOKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
DB 1 MAVMEMCPGAGSAGVQOKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60

QY 61 VITKGTAKGSEAGPAAISIIAQAECENSQBFSPFTFSERIFIAGSKQYSQESLDQIPNN 120
DB 61 VITKGTAKGSEAGPAAISIIAQAECENSQBFSPFTFSERIFIAGSKQYSQESLDQIPNN 120

QY 121 VAHATEGKMARVCKGKRKRSKARKKKKSKSLAHAGVALAKPLPTPEQESCTIPVOE 180
DB 121 VAHATEGKMARVCKGKRKRSKARKKKKSKSLAHAGVALAKPLPTPEQESCTIPVOE 180

QY 181 DESPLGAPYVNTPOFTKLPKEPGLGQLCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
DB 181 DESPLGAPYVNTPOFTKLPKEPGLGQLCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240

QY 241 KLHPDQGGPLPLTHPPFPYRLPHFPFHLQPKWPHPLPSFLGKLACVDSQKPLPDPH 300
DB 241 KLHPDQGGPLPLTHPPFPYRLPHFPFHLQPKWPHPLPSFLGKLACVDSQKPLPDPH 300

QY 301 LSKLACVDSKPLPGPHLEPCLSGAHEKSVVEYLHVALQGSVSSQASHLSLAKTW 360
DB 301 LSKLACVDSKPLPGPHLEPCLSGAHEKSVVEYLHVALQGSVSSQASHLSLAKTW 360

QY 361 AARGSRSEPSPKTDEGVLITTEKLKPDVDEYEEVHWATHQLRGLRGSGFGEVHRMEDK 420
DB 361 AARGSRSEPSPKTDEGVLITTEKLKPDVDEYEEVHWATHQLRGLRGSGFGEVHRMEDK 420

QY 421 QTGFQCAVKVRLEVFRABELMACAGLTSRIVPLYGAVREGPWVNIIFMELLEGGSLGQL 480
DB 421 QTGFQCAVKVRLEVFRABELMACAGLTSRIVPLYGAVREGPWVNIIFMELLEGGSLGQL 480

QY 481 VKEQGCLPEDRALYYLGOALGLEYLHRSRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
DB 481 VKEQGCLPEDRALYYLGOALGLEYLHRSRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540

QY 541 LQPDGLGKSLTGDYIPGTETHMAPEVVVLGRSCDAKVDWSSCCMMLHMLNGCHPWTOFF 600
DB 541 LQPDGLGKSLTGDYIPGTETHMAPEVVVLGRSCDAKVDWSSCCMMLHMLNGCHPWTOFF 600

QY 601 RGPLCLKTASPPPPVREIPSPCAPLTAQAIQGLRKEPIHRVSAEALGGKVNRLAQVGG 660
DB 601 RGPLCLKTASPPPPVREIPSPCAPLTAQAIQGLRKEPIHRVSAEALGGKVNRLAQVGG 660

QY 661 LKSPWRGKYKPRPPPPNQANYQHTLHAQRELSPRAPGPRPAETTORAPKLPPLPPE 720
DB 661 LKSPWRGKYKPRPPPPNQANYQHTLHAQRELSPRAPGPRPAETTORAPKLPPLPPE 720

QY 721 PPEPNKSPPLTLKSKEESGMWPEPLPSLLEPAPARNPSSPERKATVPPEOLOOLELEPLN 780
DB 721 PPEPNKSPPLTLKSKEESGMWPEPLPSLLEPAPARNPSSPERKATVPPEOLOOLELEPLN 780

QY 781 SLSPQFSLSEOEQILSLCLSIDSLSDSSEKNPSKASQSSRDTLSSGVHSSWSSQAEARSS 840
DB 781 SLSPQFSLSEOEQILSLCLSIDSLSDSSEKNPSKASQSSRDTLSSGVHSSWSSQAEARSS 840

QY 841 SWNVMLARGRTDTPSYFNGVKVQIQSLNGEHLHIREPHRVKVGDIATGISSQIPAAAFS 900
DB 841 SWNVMLARGRTDTPSYFNGVKVQIQSLNGEHLHIREPHRVKVGDIATGISSQIPAAAFS 900

RESULT 3
US-09-155-676A-7
; Sequence 7, Application US/09155676A
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/155.676A
; FILING DATE: 04-JAN-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IL97/00117
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117800
; FILING DATE: 02-APR-1996
; APPLICATION NUMBER: IL 119133
; FILING DATE: 26-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25, 618
; REFERENCE/DOCKET NUMBER: WALLACH-21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 947 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-155-676A-7

Query Match 100.0%; Score 5052; DB 15; Length 947;
Best Local Similarity 100.0%; Pred. No. 5.5e-306;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVMEMCPGAGSAGVQOKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
DB 1 MAVMEMCPGAGSAGVQOKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60

QY 61 VITKGTAKGSEAGPAAISIIAQAECENSQBFSPFTFSERIFIAGSKQYSQESLDQIPNN 120
DB 61 VITKGTAKGSEAGPAAISIIAQAECENSQBFSPFTFSERIFIAGSKQYSQESLDQIPNN 120

QY 121 VAHATEGKMARVCKGKRKRSKARKKKKSKSLAHAGVALAKPLPTPEQESCTIPVOE 180
DB 121 VAHATEGKMARVCKGKRKRSKARKKKKSKSLAHAGVALAKPLPTPEQESCTIPVOE 180

Db	121	VAHATEGKWARVWCKGKRSKARKKRSKSLAHAGVALAKPLPRTPEQESCTIPVQE	180
Qy	181	DESPILGAPVVRNTPOFTKPLKEPGLQGLCFKQKGELRPAALPRSELHKLISPLQCLNHVW	240
Db	181	DESPILGAPVVRNTPOFTKPLKEPGLQGLCFKQKGELRPAALPRSELHKLISPLQCLNHVW	240
Qy	241	KLHHPDQGGPLPLPHTPPPSRLPHFPFHPLOPKPHPLSFGLKACVDSQKPLPDPH	300
Db	241	KLHHPDQGGPLPLPHTPPPSRLPHFPFHPLOPKPHPLSFGLKACVDSQKPLPDPH	300
Qy	301	LKSLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW	360
Db	301	LKSLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW	360
Qy	361	AAGRSRSPSKPTEDNEGVLLTEKLPVDYREYREVEVHATHQLRGRSGFGEVHRMEDK	420
Db	361	AAGRSRSPSKPTEDNEGVLLTEKLPVDYREYREVEVHATHQLRGRSGFGEVHRMEDK	420
Qy	421	QTGFQCAVKVRLEVPRAELMACAGLTSPRIVPLYGAVREGPWNI FMELEGGSLGOL	480
Db	421	QTGFQCAVKVRLEVPRAELMACAGLTSPRIVPLYGAVREGPWNI FMELEGGSLGOL	480
Qy	481	VKEQGLCPEDRALYILGOALEGLEYLEHSRRIILHGDVKNVLLSSDGSAAALCDFGHAVC	540
Db	481	VKEQGLCPEDRALYILGOALEGLEYLEHSRRIILHGDVKNVLLSSDGSAAALCDFGHAVC	540
Qy	541	LQPDGLGKSLITGDIYIPGTETHMAPEVILGRSCDAKVDVWSSCCMMLHMLNGCHPWTQPF	600
Db	541	LQPDGLGKSLITGDIYIPGTETHMAPEVILGRSCDAKVDVWSSCCMMLHMLNGCHPWTQPF	600
Qy	601	RGPLCLKIASPPVREIPSCAPLTAQALQELRKEPIHRYSAABLGKVNRAALQOVGG	660
Db	601	RGPLCLKIASPPVREIPSCAPLTAQALQELRKEPIHRYSAABLGKVNRAALQOVGG	660
Qy	661	LKSPWRGEYKEPRHPNPQANYHTLHAOPRELSPRAPGRPAEETGSRAPKLQPLPPE	720
Db	661	LKSPWRGEYKEPRHPNPQANYHTLHAOPRELSPRAPGRPAEETGSRAPKLQPLPPE	720
Qy	721	PPENKSPPLTLKESGGMWEPIPLSLEPAPARNPSSPERKATVPEQLQLEIFLFLN	780
Db	721	PPENKSPPLTLKESGGMWEPIPLSLEPAPARNPSSPERKATVPEQLQLEIFLFLN	780
Qy	781	SLSQPFLEQEOILSLCLSDSLSDSEKNPSKASQSRDLSLSSGVHSSWSQAARS	840
Db	781	SLSQPFLEQEOILSLCLSDSLSDSEKNPSKASQSRDLSLSSGVHSSWSQAARS	840
Qy	841	SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHVRKVGDIAIGISSQIPAAAFS	900
Db	841	SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHVRKVGDIAIGISSQIPAAAFS	900
Qy	901	LVTXGQPVRYDMVPSGIDLOCTLAPDGSFAWSRVRKHQLENRP	947
Db	901	LVTXGQPVRYDMVPSGIDLOCTLAPDGSFAWSRVRKHQLENRP	947
RESULT 4			
US-09-155-676A-14			
; Sequence 14, Application US/09155676A			
; GENERAL INFORMATION:			
; APPLICANT: WALLACH, David			
; APPLICANT: MALININ, Nikolai			
; APPLICANT: BOLDIN, Mark			
; APPLICANT: KOVALENKO, Andrei			
; APPLICANT: METT, Igor			
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED			
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE			
; NUMBER OF SEQUENCES: 20			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.			
; STREET: 624 Ninth Street, N.W., Suite 300			
; CITY: Washington			
; STATE: D.C.			
; COUNTRY: USA			

; ZIP: 20001			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/155,676A			
; FILING DATE: 04-JAN-1999			
; CLASSIFICATION: 530			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: PCT/IL97/00117			
; FILING DATE: 01-APR-1997			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: IL 117800			
; FILING DATE: 02-APR-1996			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: IL 119133			
; FILING DATE: 26-AUG-1996			
; ATTORNEY/AGENT INFORMATION:			
; NAME: BROWDY, Roger L.			
; REGISTRATION NUMBER: 25, 618			
; REFERENCE/DOCKET NUMBER: WALLACH-21			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 202-628-5197			
; TELEFAX: 202-737-3528			
; INFORMATION FOR SEQ ID NO: 14:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 947 amino acids			
; TYPE: amino acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: peptide			
; US-09-155-676A-14			
Query Match 100.0%; Score 5052; DB 15; Length 947;			
Best Local Similarity 100.0%; Pred. No. 5.5e-306;			
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MAVEMACPGAGSAGVQOKELPKPKETPLPKKQSSVYKLEAVEKSPVFCGWEILND	60
Db	1	MAVEMACPGAGSAGVQOKELPKPKETPLPKKQSSVYKLEAVEKSPVFCGWEILND	60
Qy	61	VITKTAKEGSAGAPAAISIIAQACENSQSFSPFISRIIAGSKQYSQESLDOIENN	120
Db	61	VITKTAKEGSAGAPAAISIIAQACENSQSFSPFISRIIAGSKQYSQESLDOIENN	120
Qy	121	VAHATEGKWARVWCKGKRSKARKKRSKSLAHAGVALAKPLPRTPEQESCTIPVQE	180
Db	121	VAHATEGKWARVWCKGKRSKARKKRSKSLAHAGVALAKPLPRTPEQESCTIPVQE	180
Qy	181	DESPILGAPVVRNTPOFTKPLKEPGLQGLCFKQKGELRPAALPRSELHKLISPLQCLNHVW	240
Db	181	DESPILGAPVVRNTPOFTKPLKEPGLQGLCFKQKGELRPAALPRSELHKLISPLQCLNHVW	240
Qy	241	KLHHPDQGGPLPLPHTPPPSRLPHFPFHPLOPKPHPLSFGLKACVDSQKPLPDPH	300
Db	241	KLHHPDQGGPLPLPHTPPPSRLPHFPFHPLOPKPHPLSFGLKACVDSQKPLPDPH	300
Qy	301	LKSLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW	360
Db	301	LKSLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW	360
Qy	361	AAGRSRSPSKPTEDNEGVLLTEKLPVDYREYREVEVHATHQLRGRSGFGEVHRMEDK	420
Db	361	AAGRSRSPSKPTEDNEGVLLTEKLPVDYREYREVEVHATHQLRGRSGFGEVHRMEDK	420
Qy	421	QTGFQCAVKVRLEVPRAELMACAGLTSPRIVPLYGAVREGPWNI FMELEGGSLGOL	480
Db	421	QTGFQCAVKVRLEVPRAELMACAGLTSPRIVPLYGAVREGPWNI FMELEGGSLGOL	480
Qy	481	VKEQGLCPEDRALYILGOALEGLEYLEHSRRIILHGDVKNVLLSSDGSAAALCDFGHAVC	540


```
Db 1 MAVMEMACPGAPGSAGVQOKELPKPKKTPPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKGSEAGPAAISIIAAECENSQEPSTFSEIRIFTAGSKQYSQSSLDQIPNN 120
Db 61 VITKGTAKGSEAGPAAISIIAAECENSQEPSTFSEIRIFTAGSKQYSQSSLDQIPNN 120
QY 121 VAHATEGKMARVCKWKGKRSKARKKXKSSSLAHAGVALAKPLPRTPEQESCTIPVOE 180
Db 121 VAHATEGKMARVCKWKGKRSKARKKXKSSSLAHAGVALAKPLPRTPEQESCTIPVOE 180
QY 181 DESPLGAPVVRNTPQTKLPKPEGLQOLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPVVRNTPQTKLPKPEGLQOLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 240
QY 241 KLHHPDGGPLPLPHTHPFPYSRLPHFPPLQPKPHPLESFLGKLACVDSQKPLPDPH 300
Db 241 KLHHPDGGPLPLPHTHPFPYSRLPHFPPLQPKPHPLESFLGKLACVDSQKPLPDPH 300
QY 301 LSKLACVDSQKPLPGLPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSSQAHSLTSLAKTW 360
Db 301 LSKLACVDSQKPLPGLPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSSQAHSLTSLAKTW 360
QY 361 AARGSRSPSPKTEDENEGVLITKLPVDYREVEVHWATHQLRLGRSGFGEVHRMEDK 420
Db 361 AARGSRSPSPKTEDENEGVLITKLPVDYREVEVHWATHQLRLGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRABELMACAGLTSPRIPLYGAVREGPWNIIFMELLEGGSLGOL 480
Db 421 QTGFQCAVKVRLEVFRABELMACAGLTSPRIPLYGAVREGPWNIIFMELLEGGSLGOL 480
QY 481 VKEQGLCPEDRALYILGQALEGLEYLHRSRRIILHGDVKADNVLLSDGSHAALCDFGHAVC 540
Db 481 VKEQGLCPEDRALYILGQALEGLEYLHRSRRIILHGDVKADNVLLSDGSHAALCDFGHAVC 540
QY 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRCDKADVWSSCCMWLHMLNGCHPWTOFF 600
Db 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRCDKADVWSSCCMWLHMLNGCHPWTOFF 600
QY 601 RGPLCLKIASBPVPVREIIPSCAPLTAQAIQGLRKEPIHRYSAALGKGNRALQOVGG 660
Db 601 RGPLCLKIASBPVPVREIIPSCAPLTAQAIQGLRKEPIHRYSAALGKGNRALQOVGG 660
QY 661 LKSPMRGEYKEPRHPPNQANYHTLHAQPRELSRAPRPAEETTGAPKLQPLPPE 720
Db 661 LKSPMRGEYKEPRHPPNQANYHTLHAQPRELSRAPRPAEETTGAPKLQPLPPE 720
QY 721 PPEPNKSPPLTLSKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQLOLELFLN 780
Db 721 PPEPNKSPPLTLSKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQLOLELFLN 780
QY 781 SLSPQFSLBEOEQIILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHWSQAERSS 840
Db 781 SLSPQFSLBEOEQIILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHWSQAERSS 840
; TITLE OF INVENTION: Cellular Kinases Involved in Cytomegalovirus Infection and their
```

```
RESULT 7
US-09-981-397A-18
; Sequence 18, Application US/09981397A
; GENERAL INFORMATION:
; APPLICANT: Axxima Pharmaceuticals AG
; APPLICANT: Schubart, Daniel
; APPLICANT: Habenberg, Peter
; APPLICANT: Stein-Gerlach, Mathias
; APPLICANT: Bevec, Dorian
; TITLE OF INVENTION: Cellular Kinases Involved in Cytomegalovirus Infection and their
```

```
; TITLE OF INVENTION: Inhibition
; FILE REFERENCE: AXM-004.1 US
; CURRENT APPLICATION NUMBER: US/09/981,397A
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/240,750
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-981-397A-18

Query Match 100.0%; Score 5052; DB 25; Length 947;
Best Local Similarity 100.0%; Pred. No. 5.5e-306;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVMEMACPGAPGSAGVQOKELPKPKKTPPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAPGSAGVQOKELPKPKKTPPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKGSEAGPAAISIIAAECENSQEPSTFSEIRIFTAGSKQYSQSSLDQIPNN 120
Db 61 VITKGTAKGSEAGPAAISIIAAECENSQEPSTFSEIRIFTAGSKQYSQSSLDQIPNN 120
QY 121 VAHATEGKMARVCKWKGKRSKARKKXKSSSLAHAGVALAKPLPRTPEQESCTIPVOE 180
Db 121 VAHATEGKMARVCKWKGKRSKARKKXKSSSLAHAGVALAKPLPRTPEQESCTIPVOE 180
QY 181 DESPLGAPVVRNTPQTKLPKPEGLQOLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPVVRNTPQTKLPKPEGLQOLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 240
QY 241 KLHHPDGGPLPLPHTHPFPYSRLPHFPPLQPKPHPLESFLGKLACVDSQKPLPDPH 300
Db 241 KLHHPDGGPLPLPHTHPFPYSRLPHFPPLQPKPHPLESFLGKLACVDSQKPLPDPH 300
QY 301 LSKLACVDSQKPLPGLPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSSQAHSLTSLAKTW 360
Db 301 LSKLACVDSQKPLPGLPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSSQAHSLTSLAKTW 360
QY 361 AARGSRSPSPKTEDENEGVLITKLPVDYREVEVHWATHQLRLGRSGFGEVHRMEDK 420
Db 361 AARGSRSPSPKTEDENEGVLITKLPVDYREVEVHWATHQLRLGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRABELMACAGLTSPRIPLYGAVREGPWNIIFMELLEGGSLGOL 480
Db 421 QTGFQCAVKVRLEVFRABELMACAGLTSPRIPLYGAVREGPWNIIFMELLEGGSLGOL 480
QY 481 VKEQGLCPEDRALYILGQALEGLEYLHRSRRIILHGDVKADNVLLSDGSHAALCDFGHAVC 540
Db 481 VKEQGLCPEDRALYILGQALEGLEYLHRSRRIILHGDVKADNVLLSDGSHAALCDFGHAVC 540
QY 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRCDKADVWSSCCMWLHMLNGCHPWTOFF 600
Db 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRCDKADVWSSCCMWLHMLNGCHPWTOFF 600
QY 601 RGPLCLKIASBPVPVREIIPSCAPLTAQAIQGLRKEPIHRYSAALGKGNRALQOVGG 660
Db 601 RGPLCLKIASBPVPVREIIPSCAPLTAQAIQGLRKEPIHRYSAALGKGNRALQOVGG 660
QY 661 LKSPMRGEYKEPRHPPNQANYHTLHAQPRELSRAPRPAEETTGAPKLQPLPPE 720
Db 661 LKSPMRGEYKEPRHPPNQANYHTLHAQPRELSRAPRPAEETTGAPKLQPLPPE 720
QY 721 PPEPNKSPPLTLSKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQLOLELFLN 780
Db 721 PPEPNKSPPLTLSKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQLOLELFLN 780
QY 781 SLSPQFSLBEOEQIILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHWSQAERSS 840
Db 781 SLSPQFSLBEOEQIILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHWSQAERSS 840
```

```
Qy 841 SNNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSIQIPAAAFS 900
Db 841 SNNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSIQIPAAAFS 900

Qy 901 LVTKGQGVRYDMEVPSGIDLOCTLAPDGSFAWSRWKVGQLENRP 947
Db 901 LVTKGQGVRYDMEVPSGIDLOCTLAPDGSFAWSRWKVGQLENRP 947

RESULT 8
US-10-394-322A-44
; Sequence 44, Application US/10394322A
; GENERAL INFORMATION:
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.
; APPLICANT: Prescott, John C.
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS
; FILE REFERENCE: 39750-0006 US
; CURRENT APPLICATION NUMBER: US/10/394,322A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/366,892
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-394-322A-44

Query Match 100.0%; Score 5052; DB 29; Length 947;
Best Local Similarity 100.0%; Pred. No. 5.5e-306;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAVMEMACPGAGSAGVQKQELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAGSAGVQKQELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60

Qy 61 VITKGTAKGSEAGPAAISIIAQACENSQFSTFSERIFIAGSKQYSQESLDOIINN 120
Db 61 VITKGTAKGSEAGPAAISIIAQACENSQFSTFSERIFIAGSKQYSQESLDOIINN 120

Qy 121 VAHATEGMARVCKWKGRSKARKKRSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180
Db 121 VAHATEGMARVCKWKGRSKARKKRSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180

Qy 181 DESPLGAPYVRNTQFTKPLKEPGLQGLCFKQGLGEGLRPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPYVRNTQFTKPLKEPGLQGLCFKQGLGEGLRPALPRSELHKLISPLQCLNHVW 240

Qy 241 KLHPDQGGPLPLPHTPPYRSLRPPFPPLQPKWHPHLESFLGKLACVDSQKPLPDPH 300
Db 241 KLHPDQGGPLPLPHTPPYRSLRPPFPPLQPKWHPHLESFLGKLACVDSQKPLPDPH 300

Qy 301 LSKLACVDSKPLPGPHLEPCLSRGAHEKFSVEEYLVAHALQGSVSSQAHSLSLAKTW 360
Db 301 LSKLACVDSKPLPGPHLEPCLSRGAHEKFSVEEYLVAHALQGSVSSQAHSLSLAKTW 360

Qy 361 AAGRSRSPKTEDNEGVLLTEKLPVDYVEEYEEVHWATHQLRLGRSGFGEVHRMEDK 420
Db 361 AAGRSRSPKTEDNEGVLLTEKLPVDYVEEYEEVHWATHQLRLGRSGFGEVHRMEDK 420

Qy 421 QTGQCAVKVRLEVFRAEELMACGLTSPRIPLVYGAVREGPWNIIFMELLEGSLGQL 480
Db 421 QTGQCAVKVRLEVFRAEELMACGLTSPRIPLVYGAVREGPWNIIFMELLEGSLGQL 480

Qy 481 VKEQGLPEDRALYYLGALSGLEYLHSRRILHGDVADNVLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGLPEDRALYYLGALSGLEYLHSRRILHGDVADNVLSSDGSAAALCDFGHAVC 540

Qy 541 LQPDGLGKSLTGDIYIPGTETHMAPEVVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600
Db 541 LQPDGLGKSLTGDIYIPGTETHMAPEVVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600
```

```
Qy 601 RGPLCLKIASPPVREIPPSCAPLTAQAIOEGLRKEPIHRVSAAEELGGKVNRLAQVGG 660
Db 601 RGPLCLKIASPPVREIPPSCAPLTAQAIOEGLRKEPIHRVSAAEELGGKVNRLAQVGG 660

Qy 661 LKSPWGEYKEPRHPPPNQANYHOTLHAQPRELSPRAGPRPAEETTGRAKLPPLPPE 720
Db 661 LKSPWGEYKEPRHPPPNQANYHOTLHAQPRELSPRAGPRPAEETTGRAKLPPLPPE 720

Qy 721 PPEPNKSPPLTLSKEESGMWEPLPLSSLEPAPARNPSPERKATVPPELQQLIELEFLN 780
Db 721 PPEPNKSPPLTLSKEESGMWEPLPLSSLEPAPARNPSPERKATVPPELQQLIELEFLN 780

Qy 781 SLSQFFSLEEQEQLTCLSDSLSDSEKNPKASQSSRDTLSSGVHSSWSSQAERSS 840
Db 781 SLSQFFSLEEQEQLTCLSDSLSDSEKNPKASQSSRDTLSSGVHSSWSSQAERSS 840

Qy 841 SNNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSIQIPAAAFS 900
Db 841 SNNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSIQIPAAAFS 900

Qy 901 LVTKGQGVRYDMEVPSGIDLOCTLAPDGSFAWSRWKVGQLENRP 947
Db 901 LVTKGQGVRYDMEVPSGIDLOCTLAPDGSFAWSRWKVGQLENRP 947

RESULT 9
US-60-366-892-44
; Sequence 44, Application US/60366892
; GENERAL INFORMATION:
; APPLICANT: Prescott, John C.
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS
; FILE REFERENCE: SUNESIS.006PR
; CURRENT APPLICATION NUMBER: US/60/366,892
; CURRENT FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-366-892-44

Query Match 100.0%; Score 5052; DB 37; Length 947;
Best Local Similarity 100.0%; Pred. No. 5.5e-306;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAVMEMACPGAGSAGVQKQELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAGSAGVQKQELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60

Qy 61 VITKGTAKGSEAGPAAISIIAQACENSQFSTFSERIFIAGSKQYSQESLDOIINN 120
Db 61 VITKGTAKGSEAGPAAISIIAQACENSQFSTFSERIFIAGSKQYSQESLDOIINN 120

Qy 121 VAHATEGMARVCKWKGRSKARKKRSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180
Db 121 VAHATEGMARVCKWKGRSKARKKRSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180

Qy 181 DESPLGAPYVRNTQFTKPLKEPGLQGLCFKQGLGEGLRPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPYVRNTQFTKPLKEPGLQGLCFKQGLGEGLRPALPRSELHKLISPLQCLNHVW 240

Qy 241 KLHPDQGGPLPLPHTPPYRSLRPPFPPLQPKWHPHLESFLGKLACVDSQKPLPDPH 300
Db 241 KLHPDQGGPLPLPHTPPYRSLRPPFPPLQPKWHPHLESFLGKLACVDSQKPLPDPH 300

Qy 301 LSKLACVDSKPLPGPHLEPCLSRGAHEKFSVEEYLVAHALQGSVSSQAHSLSLAKTW 360
Db 301 LSKLACVDSKPLPGPHLEPCLSRGAHEKFSVEEYLVAHALQGSVSSQAHSLSLAKTW 360

Qy 361 AAGRSRSPKTEDNEGVLLTEKLPVDYVEEYEEVHWATHQLRLGRSGFGEVHRMEDK 420
Db 361 AAGRSRSPKTEDNEGVLLTEKLPVDYVEEYEEVHWATHQLRLGRSGFGEVHRMEDK 420
```



```
Db 361 AARGSRREPSPKTNEGVLLTEKLPVDYEVREBVHATHQLRGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAELMACAGLTSPRIPLYGAVREGPWVNI FWELEGGSLGOL 480
Db 421 QTGFQCAVKVRLEVFRAELMACAGLTSPRIPLYGAVREGPWVNI FWELEGGSLGOL 480
QY 481 VKEQGCLPEDRALYVLGQALEGLEYLEYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGCLPEDRALYVLGQALEGLEYLEYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGKSLLTGDYIPGTETHMAPEVILGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Db 541 LQPDGLGKSLLTGDYIPGTETHMAPEVILGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
QY 601 RGPLCLKIASBPVPVREIPSPCAPLTAQAIQEGLRKEPIHRVSAALGKGKVNRAALQOQVGG 660
Db 601 RGPLCLKIASBPVPVREIPSPCAPLTAQAIQEGLRKEPIHRVSAALGKGKVNRAALQOQVGG 660
QY 661 LKSPMRGEYKEPRHPPPNQANYHOTLHAQPRELSPPAPARNPSSPERKATVPQELQOIELEFLN 720
Db 661 LKSPMRGEYKEPRHPPPNQANYHOTLHAQPRELSPPAPARNPSSPERKATVPQELQOIELEFLN 720
QY 721 PPEPNKSPPLTLSKEESGMWEPLPLSSLEPAPARNPSSPERKATVPQELQOIELEFLN 780
Db 721 PPEPNKSPPLTLSKEESGMWEPLPLSSLEPAPARNPSSPERKATVPQELQOIELEFLN 780
QY 781 SLSQPFSLSEEQOILSCLSIDSLSDSEKPNKSKASQSSRDTLSSGVHSSSQAEARSS 840
Db 781 SLSQPFSLSEEQOILSCLSIDSLSDSEKPNKSKASQSSRDTLSSGVHSSSQAEARSS 840
QY 841 SNNMVLARGRPDTPTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
Db 841 SNNMVLARGRPDTPTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
QY 901 LVTXGQOPVRYDMEVDPDSGIDLOCTLAPDGSFAMSWRVKHGOLENRP 947
Db 901 LVTXGQOPVRYDMEVDPDSGIDLOCTLAPDGSFAMSWRVKHGOLENRP 947

RESULT 10
US-60-474-850-493
; Sequence 493, Application US/60474850
; GENERAL INFORMATION:
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Lv, Ngoc
; APPLICANT: Prentice, James
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING
; TITLE OF INVENTION: RENAL TRANSPLANT REJECTION
; FILE REFERENCE: 506613000700
; CURRENT APPLICATION NUMBER: US/60/474,850
; CURRENT FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 552
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 493
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-474-850-493

Query Match 100.0%; Score 5052; DB 37; Length 947;
Best Local Similarity 100.0%; Pred. No. 5.5e-306;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAVMEMACGAGSAGVQOQKELPKKKEKTPPLGKKOSSVYKLEAVEKSPVFCGWKEILLND 60
Db 1 MAVMEMACGAGSAGVQOQKELPKKKEKTPPLGKKOSSVYKLEAVEKSPVFCGWKEILLND 60
QY 61 VITKGTAKGSGAGPAASIIIAAECENSQEFSPFTSFERIFIAFGSKQYSQESLQIPINN 120
Db 61 VITKGTAKGSGAGPAASIIIAAECENSQEFSPFTSFERIFIAFGSKQYSQESLQIPINN 120
QY 121 VAHATEGKMARVCMWKGKRRSKARKKKSKSLAHAGVALAKPLPRTPEQBSCTIPVOE 180
```

```
Db 121 VAHATEGKMARVCMWKGKRRSKARKKKSKSLAHAGVALAKPLPRTPEQBSCTIPVOE 180
QY 181 DESPLGAPYVRNTPOFTPLKEPGLGOLCFKQLGEGLRPALPRSEIHLKILSPLOCLNHVW 240
Db 181 DESPLGAPYVRNTPOFTPLKEPGLGOLCFKQLGEGLRPALPRSEIHLKILSPLOCLNHVW 240
QY 241 KLHHPQDGGPLPLPHTPPYSRPLPQPKPPLPQPKPPLPQPKPPLPQPKPPLPQPKPPLPDPH 300
Db 241 KLHHPQDGGPLPLPHTPPYSRPLPQPKPPLPQPKPPLPQPKPPLPQPKPPLPQPKPPLPDPH 300
QY 301 LSKLACVDSKPPLPGLHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
Db 301 LSKLACVDSKPPLPGLHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
QY 361 AARGSRREPSPKTNEGVLLTEKLPVDYEVREBVHATHQLRGRSGFGEVHRMEDK 420
Db 361 AARGSRREPSPKTNEGVLLTEKLPVDYEVREBVHATHQLRGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAELMACAGLTSPRIPLYGAVREGPWVNI FWELEGGSLGOL 480
Db 421 QTGFQCAVKVRLEVFRAELMACAGLTSPRIPLYGAVREGPWVNI FWELEGGSLGOL 480
QY 481 VKEQGCLPEDRALYVLGQALEGLEYLEYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGCLPEDRALYVLGQALEGLEYLEYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGKSLLTGDYIPGTETHMAPEVILGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Db 541 LQPDGLGKSLLTGDYIPGTETHMAPEVILGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
QY 601 RGPLCLKIASBPVPVREIPSPCAPLTAQAIQEGLRKEPIHRVSAALGKGKVNRAALQOQVGG 660
Db 601 RGPLCLKIASBPVPVREIPSPCAPLTAQAIQEGLRKEPIHRVSAALGKGKVNRAALQOQVGG 660
QY 661 LKSPMRGEYKEPRHPPPNQANYHOTLHAQPRELSPPAPARNPSSPERKATVPQELQOIELEFLN 720
Db 661 LKSPMRGEYKEPRHPPPNQANYHOTLHAQPRELSPPAPARNPSSPERKATVPQELQOIELEFLN 720
QY 721 PPEPNKSPPLTLSKEESGMWEPLPLSSLEPAPARNPSSPERKATVPQELQOIELEFLN 780
Db 721 PPEPNKSPPLTLSKEESGMWEPLPLSSLEPAPARNPSSPERKATVPQELQOIELEFLN 780
QY 781 SLSQPFSLSEEQOILSCLSIDSLSDSEKPNKSKASQSSRDTLSSGVHSSSQAEARSS 840
Db 781 SLSQPFSLSEEQOILSCLSIDSLSDSEKPNKSKASQSSRDTLSSGVHSSSQAEARSS 840
QY 841 SNNMVLARGRPDTPTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
Db 841 SNNMVLARGRPDTPTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
QY 901 LVTXGQOPVRYDMEVDPDSGIDLOCTLAPDGSFAMSWRVKHGOLENRP 947
Db 901 LVTXGQOPVRYDMEVDPDSGIDLOCTLAPDGSFAMSWRVKHGOLENRP 947

RESULT 11
US-10-087-192-888
; Sequence 888, Application US/10087192
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 888
```


Db 721 PPEPNKSPPLTLTSLKEESGMWEPLPLSSLEPAPARNPSSPERKATVPEQELQLELFLN 780
Qy 781 SLSQPFSLBEQIILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHSSWSSQAEARSS 840
Db 781 SLSQPFSLBEQIILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHSSWSSQAEARSS 840
Qy 841 SNNMVLARGRPDTPDSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSISSQIPAAAFS 900
Db 841 SNNMVLARGRPDTPDSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSISSQIPAAAFS 900
Qy 901 LVTKGQPVRYDMEVPDSDGIDLQCTLAPDGSFAWSRWVKHGOLENRP 947
Db 901 LVTKGQPVRYDMEVPDSDGIDLQCTLAPDGSFAWSRWVKHGOLENRP 947

RESULT 13
US-10-170-205E-9184
; Sequence 9184, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 9184
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-9184

Query Match 99.8%; Score 5040; DB 27; Length 947;
Best Local Similarity 99.8%; Pred. No. 3.1e-305;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAVMEMACPGAGSAGVQKQKELPKPKETPPGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAGSAGVQKQKELPKPKETPPGKQSSVYKLEAVEKSPVFCGKWEILND 60

Qy 61 VITKTAGEGSBAGPAAISIIAQAECENSQEPSPTFSERIFIAGSKQYSQSSLSLQIPNN 120
Db 61 VITKTAGEGSBAGPAAISIIAQAECENSQEPSPTFSERIFIAGSKQYSQSSLSLQIPNN 120

Qy 121 VAHATEGKMARVCWKGRSKARKKRRKKSSKSLAHAGVALAKPLPRTPEQSSCTIPVQE 180
Db 121 VAHATEGKMARVCWKGRSKARKKRRKKSSKSLAHAGVALAKPLPRTPEQSSCTIPVQE 180

Qy 181 DESPLGAPVVRNTPQFTKPKLKEPGLQCLCFKOLGEGLRPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPVVRNTPQFTKPKLKEPGLQCLCFKOLGEGLRPALPRSELHKLISPLQCLNHVW 240

Qy 241 KLHHPQDGGPLPLPHTHPFYSRLPHFPFPHLPQPKPHELESFLGKLACVDSQKPLPDPH 300
Db 241 KLHHPQDGGPLPLPHTHPFYSRLPHFPFPHLPQPKPHELESFLGKLACVDSQKPLPDPH 300

Qy 301 LSKLACVDSQKPLPHTHPFYSRLPHFPFPHLPQPKPHELESFLGKLACVDSQKPLPDPH 360
Db 301 LSKLACVDSQKPLPHTHPFYSRLPHFPFPHLPQPKPHELESFLGKLACVDSQKPLPDPH 360

Qy 361 AARGSRSPSPKTEDENEGVLITLTKLPVDYREVEVHWATHQLRLGRGSFGVEVHMEDEK 420
Db 361 AARGSRSPSPKTEDENEGVLITLTKLPVDYREVEVHWATHQLRLGRGSFGVEVHMEDEK 420

Qy 421 QTGFQCAVKKVRLEVFRAELMACAGLTSPRIVPLYGAVREGPWNIFMELLEGSLGQL 480
Db 421 QTGFQCAVKKVRLEVFRAELMACAGLTSPRIVPLYGAVREGPWNIFMELLEGSLGQL 480

Qy 481 VKEQGLCPEDRALYYLGOALEGLEYLEHLSRRIILHGDVKADNVLLSSDGSNAALCDFGHAYC 540
Db 481 VKEQGLCPEDRALYYLGOALEGLEYLEHLSRRIILHGDVKADNVLLSSDGSNAALCDFGHAYC 540

Qy 541 LQPDGLGKSLTGDYIPGTETHMAPEVVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Db 541 LQPDGLGKSLTGDYIPGTETHMAPEVVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600

Qy 601 RGPLCLKIASEPPPVREIIPSCAPLTAQAIQEGLRKEPIHHRVSAAEELGKVNRAALQOQVG 660
Db 601 RGPLCLKIASEPPPVREIIPSCAPLTAQAIQEGLRKEPIHHRVSAAEELGKVNRAALQOQVG 660

Qy 661 LKSPMRGKYKPRHPPPNQANYHOTLHAQPRELSPRAPGPRPAEETTCRAPKLQPLPPE 720
Db 661 LKSPMRGKYKPRHPPPNQANYHOTLHAQPRELSPRAPGPRPAEETTCRAPKLQPLPPE 720

Qy 721 PPEPNKSPPLTLTSLKEESGMWEPLPLSSLEPAPARNPSSPERKATVPEQELQLELFLN 780
Db 721 PPEPNKSPPLTLTSLKEESGMWEPLPLSSLEPAPARNPSSPERKATVPEQELQLELFLN 780

Qy 781 SLSQPFSLBEQIILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHSSWSSQAEARSS 840
Db 781 SLSQPFSLBEQIILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHSSWSSQAEARSS 840

Qy 841 SNNMVLARGRPDTPDSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSISSQIPAAAFS 900
Db 841 SNNMVLARGRPDTPDSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSISSQIPAAAFS 900

Qy 901 LVTKGQPVRYDMEVPDSDGIDLQCTLAPDGSFAWSRWVKHGOLENRP 947
Db 901 LVTKGQPVRYDMEVPDSDGIDLQCTLAPDGSFAWSRWVKHGOLENRP 947

RESULT 14
US-60-443-566-2590
; Sequence 2590, Application US/60443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001447
; CURRENT APPLICATION NUMBER: US/60/443,566
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2590
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-443-566-2590

Query Match 99.8%; Score 5040; DB 37; Length 947;
Best Local Similarity 99.8%; Pred. No. 3.1e-305;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAVMEMACPGAGSAGVQKQKELPKPKETPPGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAGSAGVQKQKELPKPKETPPGKQSSVYKLEAVEKSPVFCGKWEILND 60

Qy 61 VITKTAGEGSBAGPAAISIIAQAECENSQEPSPTFSERIFIAGSKQYSQSSLSLQIPNN 120
Db 61 VITKTAGEGSBAGPAAISIIAQAECENSQEPSPTFSERIFIAGSKQYSQSSLSLQIPNN 120

Qy 121 VAHATEGKMARVCWKGRSKARKKRRKKSSKSLAHAGVALAKPLPRTPEQSSCTIPVQE 180
Db 121 VAHATEGKMARVCWKGRSKARKKRRKKSSKSLAHAGVALAKPLPRTPEQSSCTIPVQE 180

Qy 181 DESPLGAPVVRNTPQFTKPKLKEPGLQCLCFKOLGEGLRPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPVVRNTPQFTKPKLKEPGLQCLCFKOLGEGLRPALPRSELHKLISPLQCLNHVW 240

Qy 241 KLHHPQDGGPLPLPHTHPFYSRLPHFPFPHLPQPKPHELESFLGKLACVDSQKPLPDPH 300
Db 241 KLHHPQDGGPLPLPHTHPFYSRLPHFPFPHLPQPKPHELESFLGKLACVDSQKPLPDPH 300

Qy 301 LSKLACVDSQKPLPHTHPFYSRLPHFPFPHLPQPKPHELESFLGKLACVDSQKPLPDPH 360
Db 301 LSKLACVDSQKPLPHTHPFYSRLPHFPFPHLPQPKPHELESFLGKLACVDSQKPLPDPH 360

```
Db 301 LSKLACVDSKPPLPGPHLEPSCLSGAHEKSVSEYLVHALQGSVSSQAHSLSLAKTW 360
Qy 361 AARGSRSEPPKPTEDNEGVLITTEKLKPDVDEYEEVHWAHQRLGRGSGFGEVHRMEDK 420
Db 361 AARGSRSEPPKPTEDNEGVLITTEKLKPDVDEYEEVHWAHQRLGRGSGFGEVHRMEDK 420
Qy 421 QTGFQCAVKKVRLEVFRAEELMACAGLTSPRIVLYGAVREGPWNIIFMELLEGGSLGOL 480
Db 421 QTGFQCAVKKVRLEVFRAEELMACAGLTSPRIVLYGAVREGPWNIIFMELLEGGSLGOL 480
Qy 481 VKEQCLPEDRALYVYLGQALEGLVYHSHRRILHGDVADNVLSSDGSAAALCDFGHAVC 540
Db 481 VKEQCLPEDRALYVYLGQALEGLVYHSHRRILHGDVADNVLSSDGSAAALCDFGHAVC 540
Qy 541 LQPDGLGKSLITGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Db 541 LQPDGLGKSLITGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Qy 601 RGPLCLKIASPPPPVREIPSPCAPLTAQAIQEGLRKEPIHRVSAELGGKVNRAALQVGG 660
Db 601 RGPLCLKIASPPPPVREIPSPCAPLTAQAIQEGLRKEPIHRVSAELGGKVNRAALQVGG 660
Qy 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPAPBETTTGRAPKLOPPLPPE 720
Db 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPAPBETTTGRAPKLOPPLPPE 720
Qy 721 PPEPNKSPPLTLKSKEESGMWEPPLSSLEPAPARNPSSPERKATVPQEOLQOLEIEFLN 780
Db 721 PPEPNKSPPLTLKSKEESGMWEPPLSSLEPAPARNPSSPERKATVPQEOLQOLEIEFLN 780
Qy 781 SLSQPFSLSEEQEQLISCLSIDSLSDSDSEKNPKSKASQSSRDTLSSGVHSSQAARSS 840
Db 781 SLSQPFSLSEEQEQLISCLSIDSLSDSDSEKNPKSKASQSSRDTLSSGVHSSQAARSS 840
Qy 841 SNNVLAGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAIGISSQIPAAAFS 900
Db 841 SNNVLAGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAIGISSQIPAAAFS 900
Qy 901 LVTKDQGPVRYDMEVPDSDGLDQCTLAPDGSFAMSWRVKXGQLENRP 947
Db 901 LVTKDQGPVRYDMEVPDSDGLDQCTLAPDGSFAMSWRVKXGQLENRP 947

RESULT 15
US-60-443-566-2591
; Sequence 2591, Application US/60443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001447
; CURRENT APPLICATION NUMBER: US/60/443,566
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2591
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-443-566-2591

Query Match 99.8%; Score 5040; DB 37; Length 947;
Best Local Similarity 99.8%; Pred.No. 3.1e-305;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAVMEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEIIND 60
Db 1 MAVMEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEIIND 60
Qy 61 VITGTAKESGAGPAAISIIAQAECENSQFSTFSEIRIFIAQSKQYSQESLDOIINN 120
```

Search completed: June 7, 2005, 12:19:39
Job time : 427.493 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:51:20 ; Search time 37.0523 Seconds
(without alignments)
1907.914 Million cell updates/sec

Title: US-09-155-676B-7
Perfect score: 5052
Sequence: 1 MAVMEMACPGAPGSAVGQOK.....PDGSFAMSWRVKHQLENRP 947

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pcp:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pcp:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pcp:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pcp:*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pcp:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5052	100.0	947	3	US-09-257-703-1
2	5052	100.0	947	4	US-09-871-889A-1
3	5044	99.8	947	2	US-08-887-518-2
4	5044	99.8	947	2	US-09-023-321-2
5	5044	99.8	947	2	US-09-032-475-2
6	479	9.5	94	4	US-09-513-999C-6660
7	360.5	7.1	467	4	US-09-522-775A-4
8	360	7.1	467	4	US-09-522-775A-2
9	360	7.1	475	4	US-09-949-016-10456
10	358.5	7.1	1302	3	US-09-423-890-2
11	355.5	7.0	672	1	US-08-049-254-2
12	355.5	7.0	672	1	US-08-472-934-2
13	355.5	7.0	672	2	US-08-323-460A-2
14	355.5	7.0	672	2	US-08-461-146C-2
15	355.5	7.0	672	3	US-08-461-145C-2
16	355.5	7.0	672	3	US-08-628-829-2
17	355.5	7.0	1492	4	US-09-697-898-5
18	355.5	7.0	1493	3	US-09-423-890-8
19	355.5	7.0	1593	3	US-08-628-829-4
20	355	7.0	1512	4	US-09-697-898-2
21	353	7.0	619	1	US-08-472-934-4
22	353	7.0	619	2	US-08-323-460A-4
23	353	7.0	619	3	US-08-461-146C-4
24	353	7.0	619	3	US-08-461-145C-4
25	353	7.0	619	3	US-09-423-890-10
26	353	7.0	619	3	US-08-628-829-6
27	353	7.0	1492	4	US-09-697-898-4

28	352	7.0	619	1	US-08-472-934-12	Sequence 12, Appl
29	352	7.0	619	2	US-08-461-146C-12	Sequence 12, Appl
30	352	7.0	619	3	US-08-461-145C-12	Sequence 12, Appl
31	352	7.0	619	3	US-08-628-829-8	Sequence 8, Appl
32	348	6.9	619	3	US-09-423-890-4	Sequence 4, Appl
33	346	6.8	1495	4	US-09-697-898-3	Sequence 3, Appl
34	326.5	6.5	651	4	US-09-371-338-15	Sequence 15, Appl
35	322	6.4	268	4	US-09-371-338-17	Sequence 17, Appl
36	322	6.4	647	3	US-09-031-563-7	Sequence 7, Appl
37	322	6.4	647	4	US-09-392-277-7	Sequence 7, Appl
38	322	6.4	647	4	US-09-258-000-7	Sequence 7, Appl
39	322	6.4	648	3	US-09-031-563-5	Sequence 5, Appl
40	322	6.4	648	4	US-09-392-277-5	Sequence 5, Appl
41	322	6.4	648	4	US-09-258-000-5	Sequence 5, Appl
42	322	6.4	1315	3	US-09-031-563-2	Sequence 2, Appl
43	322	6.4	1315	3	US-09-293-505-10	Sequence 10, Appl
44	322	6.4	1315	4	US-09-392-277-2	Sequence 2, Appl
45	322	6.4	1315	4	US-09-258-000-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-09-257-703-1
; Sequence 1, Application US/09257703
; Patent No. 6265538
; GENERAL INFORMATION:
; APPLICANT: Greene, Warner C.
; APPLICANT: Lin, Xin
; APPLICANT: Gelezunas, Romas
; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED
; FILE REFERENCE: 30448.61USU1
; CURRENT APPLICATION NUMBER: US/09/257,703
; CURRENT FILING DATE: 1999-02-25
; EARLIER APPLICATION NUMBER: 60/076,299
; EARLIER FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 947
; TYPE: PRT
; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)
US-09-257-703-1

Query Match	100.0%	Score	5052;	DB	3;	Length	947;
Best Local Similarity	100.0%;	Pred. No.	0;	Mismatches	0;	Indels	0;
Matches	947;	Conservative	0;				
Qy	1	MAVMEMACPGAPGSAVGQOKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND	60				
Db	1	MAVMEMACPGAPGSAVGQOKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND	60				
Qy	61	VITKGTAGESEAGPAAISIIAQACENSQSFPTFSRIFIAQSGSKSLAHAGVALAKPLPRTPEQESCTIPVOE	120				
Db	61	VITKGTAGESEAGPAAISIIAQACENSQSFPTFSRIFIAQSGSKSLAHAGVALAKPLPRTPEQESCTIPVOE	120				
Qy	121	VAHATEGMARVCKGKRSKARKKKKSSKSLAHAGVALAKPLPRTPEQESCTIPVOE	180				
Db	121	VAHATEGMARVCKGKRSKARKKKKSSKSLAHAGVALAKPLPRTPEQESCTIPVOE	180				
Qy	181	DESLPLGAPYVNTQFTKPLKEPGLQCFKQLEGRLPALPRSELHKLISPLQCLNHVW	240				
Db	181	DESLPLGAPYVNTQFTKPLKEPGLQCFKQLEGRLPALPRSELHKLISPLQCLNHVW	240				
Qy	241	KLHPQDGGPLPLTHPPFPYSLRPLPFPHPQPKPPLPSFLGKLACVDSQKPLPDHP	300				
Db	241	KLHPQDGGPLPLTHPPFPYSLRPLPFPHPQPKPPLPSFLGKLACVDSQKPLPDHP	300				
Qy	301	LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVAHQGVSSSQASHLSLAKTW	360				
Db	301	LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVAHQGVSSSQASHLSLAKTW	360				

361 AARGSRSEPSKTEDNEGVLITKLPVDYREVEVHATHOLRLGRSGFEVHRMEDK 420
Db |||||
361 AARGSRSEPSKTEDNEGVLITKLPVDYREVEVHATHOLRLGRSGFEVHRMEDK 420
Qy |||||
421 QTGFQCAVKKVLEVFRAEELMACAGLTSPRI VPLYGAVREGPWNI FMELEGGSLGQL 480
Db |||||
421 QTGFQCAVKKVLEVFRAEELMACAGLTSPRI VPLYGAVREGPWNI FMELEGGSLGQL 480
Qy |||||
481 VKEQCLPEDRALYVIGQALEGLYLHSGRRLIHGVDKADNVLSSDGSAAALCDFGHAVC 540
Db |||||
481 VKEQCLPEDRALYVIGQALEGLYLHSGRRLIHGVDKADNVLSSDGSAAALCDFGHAVC 540
Qy |||||
541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Db |||||
541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Qy |||||
601 RGPLCLKIASEPPPVREIIPSCAPLTAQAI QEGLRKEPIHRSAAELGGKVNRAALQQVGG 660
Db |||||
601 RGPLCLKIASEPPPVREIIPSCAPLTAQAI QEGLRKEPIHRSAAELGGKVNRAALQQVGG 660
Qy |||||
661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPRPAEETTGRAPKLQPLPPE 720
Db |||||
661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPRPAEETTGRAPKLQPLPPE 720
Qy |||||
721 PPEPNKSPPLTSLKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQLELFLN 780
Db |||||
721 PPEPNKSPPLTSLKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQLELFLN 780
Qy |||||
781 SLSQPFSLBEEQILSCLSDSLSDSEKKNPSKASQSSRDTLSSGVHWSQAARSS 840
Db |||||
781 SLSQPFSLBEEQILSCLSDSLSDSEKKNPSKASQSSRDTLSSGVHWSQAARSS 840
Qy |||||
841 SWNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
Db |||||
841 SWNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
Qy |||||
901 LVTKDQGPVRYDMEVPDSDGIDLQCTLAPDGSFAMSWRVKHGQLENRP 947
Db |||||
901 LVTKDQGPVRYDMEVPDSDGIDLQCTLAPDGSFAMSWRVKHGQLENRP 947
Qy |||||

RESULT 2
US-09-871-889A-1
; Sequence 1, Application US/09871889A
; Patent No. 6645728
; GENERAL INFORMATION:
; APPLICANT: Greene, Warner C.
; APPLICANT: Lin, Xin
; APPLICANT: Gelezuinas, Romas
; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED BY TNF-ALPHA
; FILE REFERENCE: 30448.61USD1
; CURRENT APPLICATION NUMBER: US/09/871.889A
; PRIOR FILING DATE: 2001-06-01
; PRIOR FILING DATE: 1999-02-25
; PRIOR FILING DATE: 1999-02-25
; PRIOR FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 947
; TYPE: PRT
; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)
US-09-871-889A-1

Query Match 100.0%; Score 5052; DB 4; Length 947;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAVMEMACPGAGSVAVGQKELPKPKETPLPGKQSSVYKLEAVEKSPVFCGWEILND 60
|||

1 MAVMEMACPGAGSVAVGQKELPKPKETPLPGKQSSVYKLEAVEKSPVFCGWEILND 60
Db |||||
61 VITKGTAKESGAPAAISIIIAQECENSQEFPTSERIFIAGSKQYSQESLQIPNN 120
Qy |||||
61 VITKGTAKESGAPAAISIIIAQECENSQEFPTSERIFIAGSKQYSQESLQIPNN 120
Db |||||
121 VAHATEGMARVCKWGRKRSKARKKRKKSLSLAHAGVALAKPLPRTPEQESCTIPVQE 180
Qy |||||
121 VAHATEGMARVCKWGRKRSKARKKRKKSLSLAHAGVALAKPLPRTPEQESCTIPVQE 180
Db |||||
181 DESPLGAPYVRNTPTFTKLPKEGLGOI.CPKQLGEGRLPALPRSELHKLISPLQCLNHVW 240
Qy |||||
181 DESPLGAPYVRNTPTFTKLPKEGLGOI.CPKQLGEGRLPALPRSELHKLISPLQCLNHVW 240
Db |||||
241 KLHHPQDGGPLPLPPTHPPYPSRLPHPPHPLQPKWHPLESFLGKLACVDSQKPLDPH 300
Qy |||||
241 KLHHPQDGGPLPLPPTHPPYPSRLPHPPHPLQPKWHPLESFLGKLACVDSQKPLDPH 300
Db |||||
301 LSKLACVDSQKPLPFGPHLEPSCLSRGAHEKFSVEEYLVHALQGVSSQAHSLSLAKTW 360
Qy |||||
301 LSKLACVDSQKPLPFGPHLEPSCLSRGAHEKFSVEEYLVHALQGVSSQAHSLSLAKTW 360
Db |||||
361 AARGSRSEPSKTEDNEGVLITKLPVDYREVEVHATHOLRLGRSGFEVHRMEDK 420
Qy |||||
361 AARGSRSEPSKTEDNEGVLITKLPVDYREVEVHATHOLRLGRSGFEVHRMEDK 420
Db |||||
421 QTGFQCAVKKVLEVFRAEELMACAGLTSPRI VPLYGAVREGPWNI FMELEGGSLGQL 480
Qy |||||
421 QTGFQCAVKKVLEVFRAEELMACAGLTSPRI VPLYGAVREGPWNI FMELEGGSLGQL 480
Db |||||
481 VKEQCLPEDRALYVIGQALEGLYLHSGRRLIHGVDKADNVLSSDGSAAALCDFGHAVC 540
Qy |||||
481 VKEQCLPEDRALYVIGQALEGLYLHSGRRLIHGVDKADNVLSSDGSAAALCDFGHAVC 540
Db |||||
541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Qy |||||
541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Db |||||
601 RGPLCLKIASEPPPVREIIPSCAPLTAQAI QEGLRKEPIHRSAAELGGKVNRAALQQVGG 660
Qy |||||
601 RGPLCLKIASEPPPVREIIPSCAPLTAQAI QEGLRKEPIHRSAAELGGKVNRAALQQVGG 660
Db |||||
661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPRPAEETTGRAPKLQPLPPE 720
Qy |||||
661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPRPAEETTGRAPKLQPLPPE 720
Db |||||
721 PPEPNKSPPLTSLKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQLELFLN 780
Qy |||||
721 PPEPNKSPPLTSLKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQLELFLN 780
Db |||||
781 SLSQPFSLBEEQILSCLSDSLSDSEKKNPSKASQSSRDTLSSGVHWSQAARSS 840
Qy |||||
781 SLSQPFSLBEEQILSCLSDSLSDSEKKNPSKASQSSRDTLSSGVHWSQAARSS 840
Db |||||
841 SWNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
Qy |||||
841 SWNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
Db |||||
901 LVTKDQGPVRYDMEVPDSDGIDLQCTLAPDGSFAMSWRVKHGQLENRP 947
Qy |||||
901 LVTKDQGPVRYDMEVPDSDGIDLQCTLAPDGSFAMSWRVKHGQLENRP 947
Db |||||

RESULT 3
US-08-887-518-2
; Sequence 2, Application US/08887518
; Patent No. 5843721
; GENERAL INFORMATION:
; APPLICANT: Rothe, Mike
; APPLICANT: Wu, Lin
; TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
;

Query Match	99.8%	Score 5044;	DB 2;	Length 947;
Best Local Similarity	99.9%	Pred. No. 0;		

Matches	946;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	MVMEACPGAPGSAVGOQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND	60						
Db	1	MVMEACPGAPGSAVGOQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND	60						
QY	61	VITKGTAKGSEAGPAAISIIAQAECENSQEFSPFTSERIFIAGSKQYSQSESLDQIPNN	120						
Db	61	VITKGTAKGSEAGPAAISIIAQAECENSQEFSPFTSERIFIAGSKQYSQSESLDQIPNN	120						
QY	121	VAHATEGMARVCWKGRSKARKKRRKSKSLAHAGVALAKPLRTPPEQESCTIPVOE	180						
Db	121	VAHATEGMARVCWKGRSKARKKRRKSKSLAHAGVALAKPLRTPPEQESCTIPVOE	180						
QY	181	DESPLAGAPYVRNTPQFTKPLKEPGLQCFKQGLGEGRLPALPSELHKLISPLQCLNHVW	240						
Db	181	DESPLAGAPYVRNTPQFTKPLKEPGLQCFKQGLGEGRLPALPSELHKLISPLQCLNHVW	240						
QY	241	KLHPQDGGPLPLTHPPYSRPLPFPHPPLQPKPKHPLESPLGKLACVDSQKPLDPDH	300						
Db	241	KLHPQDGGPLPLTHPPYSRPLPFPHPPLQPKPKHPLESPLGKLACVDSQKPLDPDH	300						
QY	301	LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLTSLAKTW	360						
Db	301	LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLTSLAKTW	360						
QY	361	AAGRSRSPSPKPTEDNEGVLITKLPVDYEVREEVHWATHQLRGRGSGFGEVHRMEDK	420						
Db	361	AAGRSRSPSPKPTEDNEGVLITKLPVDYEVREEVHWATHQLRGRGSGFGEVHRMEDK	420						
QY	421	QTGFQCAVKKRLVLEVFRAELMACAGLTSPRIPLYGAVREGPWVNFMELEGGSLGQL	480						
Db	421	QTGFQCAVKKRLVLEVFRAELMACAGLTSPRIPLYGAVREGPWVNFMELEGGSLGQL	480						
QY	481	VKEQGLCPEDRALYILGQALEGLYLEHSRIILHGDVKADNVLSSDGSAAALCDFGHAVC	540						
Db	481	VKEQGLCPEDRALYILGQALEGLYLEHSRIILHGDVKADNVLSSDGSAAALCDFGHAVC	540						
QY	541	LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMWMLHMLNGCHPWTOFF	600						
Db	541	LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMWMLHMLNGCHPWTOFF	600						
QY	601	RGPLCLKIASPPPPVREIPSPCAPLTAQAIQEGRLKEPIHRVSAALGKGKVNRLQQVGG	660						
Db	601	RGPLCLKIASPPPPVREIPSPCAPLTAQAIQEGRLKEPIHRVSAALGKGKVNRLQQVGG	660						
QY	661	LKSPWGEYKEPRHPPPNQAANYHOTLHAQPRELSRAPGPRPAEATTGAPKLPPLPPE	720						
Db	661	LKSPWGEYKEPRHPPPNQAANYHOTLHAQPRELSRAPGPRPAEATTGAPKLPPLPPE	720						
QY	721	PPEPNKSPPLTLSKESGMEWELPLSSLEPAPARNPSPERKATVPPELOLELEFLN	780						
Db	721	PPEPNKSPPLTLSKESGMEWELPLSSLEPAPARNPSPERKATVPPELOLELEFLN	780						
QY	781	SLSQPFLBEQOILSLCLSIDSLSDSEKNPKSKASSRDLTSSGVHSSWSSQAEARSS	840						
Db	781	SLSQPFLBEQOILSLCLSIDSLSDSEKNPKSKASSRDLTSSGVHSSWSSQAEARSS	840						
QY	841	SNMVLARGRPDTPSYFNGVKVQIQSLNGEHLHIREHRVKVGDIAATGSISSQIPAAAFS	900						
Db	841	SNMVLARGRPDTPSYFNGVKVQIQSLNGEHLHIREHRVKVGDIAATGSISSQIPAAAFS	900						
QY	901	LVTKDQQPVRYDMEVPDSDGIDLCQTLAPDGSFAMSWRVKHGQLENRP	947						
Db	901	LVTKDQQPVRYDMEVPDSDGIDLCQTLAPDGSFAMSWRVKHGQLENRP	947						

RESULT 5
US-09-032-475-2
; Sequence 2, Application US/09032475
; Patent No. 5854003
; GENERAL INFORMATION:
; APPLICANT: Roche, Mike

APPLICANT: Wu, Lin
TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032.475
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/887,518
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: T97-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 947 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-032-475-2

Query Match 99.8%; Score 5044; DB 2; Length 947;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 946; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MVMEACPGAPGSAVGOQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND	60
Db	1	MVMEACPGAPGSAVGOQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND	60
QY	61	VITKGTAKGSEAGPAAISIIAQAECENSQEFSPFTSERIFIAGSKQYSQSESLDQIPNN	120
Db	61	VITKGTAKGSEAGPAAISIIAQAECENSQEFSPFTSERIFIAGSKQYSQSESLDQIPNN	120
QY	121	VAHATEGMARVCWKGRSKARKKRRKSKSLAHAGVALAKPLRTPPEQESCTIPVOE	180
Db	121	VAHATEGMARVCWKGRSKARKKRRKSKSLAHAGVALAKPLRTPPEQESCTIPVOE	180
QY	181	DESPLAGAPYVRNTPQFTKPLKEPGLQCFKQGLGEGRLPALPSELHKLISPLQCLNHVW	240
Db	181	DESPLAGAPYVRNTPQFTKPLKEPGLQCFKQGLGEGRLPALPSELHKLISPLQCLNHVW	240
QY	241	KLHPQDGGPLPLTHPPYSRPLPFPHPPLQPKPKHPLESPLGKLACVDSQKPLDPDH	300
Db	241	KLHPQDGGPLPLTHPPYSRPLPFPHPPLQPKPKHPLESPLGKLACVDSQKPLDPDH	300
QY	301	LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLTSLAKTW	360
Db	301	LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLTSLAKTW	360
QY	361	AAGRSRSPSPKPTEDNEGVLITKLPVDYEVREEVHWATHQLRGRGSGFGEVHRMEDK	420
Db	361	AAGRSRSPSPKPTEDNEGVLITKLPVDYEVREEVHWATHQLRGRGSGFGEVHRMEDK	420
QY	421	QTGFQCAVKKRLVLEVFRAELMACAGLTSPRIPLYGAVREGPWVNFMELEGGSLGQL	480
Db	421	QTGFQCAVKKRLVLEVFRAELMACAGLTSPRIPLYGAVREGPWVNFMELEGGSLGQL	480

QY 481 VKQCCLPEDRALYYLGOALEGLEYLHSHRRILHGDVADNVLSSDGSAAALCDPFGHVC 540
Db 481 VKQCCLPEDRALYYLGOALEGLEYLHSHRRILHGDVADNVLSSDGSAAALCDPFGHVC 540
QY 541 LOPDGLGSLTGDYIPCTETHMAPEVVLGRSCDAKVDVWSSCCMLHMLNGCHPWTOFF 600
Db 541 LOPDGLGSLTGDYIPCTETHMAPEVVLGRSCDAKVDVWSSCCMLHMLNGCHPWTOFF 600
QY 601 RGPLCLKIASEPPPPVREIPPPSCAPLTAQAIQEGRLKEPIHRVSAELGKYNRALQQVGG 660
Db 601 RGPLCLKIASEPPPPVREIPPPSCAPLTAQAIQEGRLKEPIHRVSAELGKYNRALQQVGG 660
QY 661 LKSPWRGKYKPRHPPPNQANHYOTLHAQPRELSPRAPGPRPAETTTGRAPKLQPPLPPE 720
Db 661 LKSPWRGKYKPRHPPPNQANHYOTLHAQPRELSPRAPGPRPAETTTGRAPKLQPPLPPE 720
QY 721 PPEPNKSPPLTLKSEESGMBEPLPLSSLEPAPARNPSSPERKATVPEQELQQLLELFLN 780
Db 721 PPEPNKSPPLTLKSEESGMBEPLPLSSLEPAPARNPSSPERKATVPEQELQQLLELFLN 780
QY 781 SLSQPFSLSEEOQLISCLISLSDSDSEKNPKSKASOSSRDTLSSGVHSSQAEARSS 840
Db 781 SLSQPFSLSEEOQLISCLISLSDSDSEKNPKSKASOSSRDTLSSGVHSSQAEARSS 840
QY 841 SNNVYLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREPHRVKVGDIATGISSQIPAAAFS 900
Db 841 SNNVYLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREPHRVKVGDIATGISSQIPAAAFS 900
QY 901 LVTKDQGVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKHQLENRP 947
Db 901 LVTKDQGVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKHQLENRP 947

RESULT 6
US-09-513-999C-6660
; Sequence 6660, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59. US2. REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122.487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6660
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-6660

Query Match 9.5%; Score 479; DB 4; Length 94;
Best Local Similarity 98.9%; Pred. No. 6.4e-29;
Matches 93; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAVMEMCPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMCPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKGSEAGAPAAISIIAQAECENSQEFSP 94
Db 61 VITKGTAKGSEAGAPAAISIIAQAECENSQEFSP 94

RESULT 7
US-09-522-775A-4
; Sequence 4, Application US/09522775A

; Patent No. 6660906
; GENERAL INFORMATION:
; APPLICANT: Tsichlis, Philip N.
; TITLE OF INVENTION: Inhibition of Tpl2 To Treat Inflammatory Diseases
; FILE REFERENCE: OTT-3202
; CURRENT APPLICATION NUMBER: US/09/522.775A
; CURRENT FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Rat Tpl2 polypeptide sequence
US-09-522-775A-4
Query Match 7.1%; Score 360.5; DB 4; Length 467;
Best Local Similarity 27.9%; Pred. No. 5.3e-19;
Matches 107; Conservative 66; Mismatches 127; Indels 83; Gaps 13;
QY 332 SVEEYLVHALQGSVSSQASHLSLAKTWAARGSRSPKTEDNEGVLLEKLPVDY 391
Db 80 TVEDLL--AFANHISNTKHFY-----GCRPQE-----SGILLNMWISPOG 119
QY 392 EYREE-----VHWATHOLRLG-----RGSFGEVHRMEDKOTGFCQCAVKKVLEVFRAEEL 441
Db 120 RYQIDSDVLLVPWKLTYSISGSGFVPRGAFKVLQAQDMKTKRMACKLIPVDQFKPSDV 179
QY 442 MACAGLTSPIVPLYGAVREGPWNIWELLEGSLQGVKEQGLPDRALYYLGOALE 501
Db 180 EIQACFRHENIAELYGAVLWMDTVHLEFMEAGEGSLVLEKLESCGPMREFEIIVTWKHLVK 239
QY 502 GLEYLHSHRRILHGDVADNVLSSDGSAAALCDPFGHVCVCLQPDGLGKSLLTGDYIP---- 557
Db 240 GLDFLHKKKVIHDKIPSNIVFMS--TKAVLVDFGLSVQMTED-----VTLPKDLR 288
QY 558 GTETHMAPEVVLGRSCDAKVDVWSSCCMLHMLNGCHPWTOFF---RGPLCLKIA-SEPP 613
Db 289 GTEIYMSPEVILCRGHSTKADIYSLGATLIHMOTGTPPWVKRYPRSAVPSYLYIHKQAP 348
QY 614 PVRIIPSPCAPLTAQAIQEGRLKEPIHRVSAELGKYNRALQQVGGGLKSPWGEYKEPR 673
Db 349 PLEDIAGDCSPGMRLEIEAALERNPNRPKAADL-----LK----- 384
QY 674 HPPNQANYHOTLHAQPRELSR 696
Db 385 -----HEALN-PPREDQPR 397

RESULT 8
US-09-522-775A-2
; Sequence 2, Application US/09522775A
; Patent No. 6660906
; GENERAL INFORMATION:
; APPLICANT: Tsichlis, Philip N.
; TITLE OF INVENTION: Inhibition of Tpl2 To Treat Inflammatory Diseases
; FILE REFERENCE: OTT-3202
; CURRENT APPLICATION NUMBER: US/09/522.775A
; CURRENT FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human Tpl2 polypeptide sequence
US-09-522-775A-2

Query Match 7.1%; Score 360; DB 4; Length 467;
Best Local Similarity 27.2%; Pred. No. 5.8e-19;
Matches 116; Conservative 69; Mismatches 155; Indels 86; Gaps 14;

QY 332 SVEEYLVAHQSGVSSOAHSLTSLAKTWAAGRSRREPSPKTNEGVLLTEKLKVDY 391
Db 80 TVEDLL--AFANHISNTAKHY-----GQRPQE-----SGILLNNVITPONG 119
QY 392 EYREE-----VHWATHQLRLG-----RGSFGVHRMEDKQTGFQCAVKKVRLEVPRAEEL 441
Db 120 RYQIDSDVLLIPWKLTYRNIGSDIFPRGAFGVYLAQDIKTKRMACKLIPVDQFKPSDV 179
QY 442 MACAGLTSPIRVLYGAVREGPWNIEMELLEGGSLGQLVKEOGCLPEDRALYYLQGALE 501
Db 180 EIQAQCFRHENIAELYGAVLWGETVHLFMEAGEGSGVLEKLESCGPMREFEIIVWTKHVLK 239
QY 502 GLEYLHRSRILHGDVADNVLLSSDGSAAALCDFGHAVCLQDPDGLKSLLTGDYIP---- 557
Db 240 GLDFLHSHKVIHHDIKPSNIVFMS--TKAVLVDVFGLSVQMTED-----VYFPKDLR 288
QY 558 GTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF---RGPLCLKIA-SEPP 613
Db 289 GTEIYMSPEVILCRGHSTKADIYSLGATLIHMOTGTPPWVKRYPRSAVPSYLYIIHKQAP 348
QY 614 PVREIIPSCAPLTAQAOIEGLRKEPIHRVSAALGGKVNRAALQQVGGGLKSPWRGEYKEPR 673
Db 349 PLEDIADDCSPGMRLEIEASLERPNHPRADL-----LK----- 384
QY 674 HPPPNQANYHOTLHAQPRELSPPRAGPRAEETTGRAKLPQPLPPEPPENKSPPLTUS 733
Db 385 -----HEALN-PPREDQPRC---QSLDSALLERKLLSRKELELPENIADSSCTGS 431
QY 734 KEESGM 739
Db 432 TEESEM 437
RESULT 9
US-09-949-016-10456
; Sequence 10456, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J, Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10456
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10456
Query Match 7.1%; Score 360; DB 4; Length 475;
Best Local Similarity 27.2%; Pred. No. 5.9e-19;
Matches 116; Conservative 69; Mismatches 155; Indels 86; Gaps 14;
QY 332 SVEEYLVAHQSGVSSOAHSLTSLAKTWAAGRSRREPSPKTNEGVLLTEKLKVDY 391
Db 88 TVEDLL--AFANHISNTAKHY-----GQRPQE-----SGILLNNVITPONG 127
QY 392 EYREE-----VHWATHQLRLG-----RGSFGVHRMEDKQTGFQCAVKKVRLEVPRAEEL 441
Db 128 RYQIDSDVLLIPWKLTYRNIGSDIFPRGAFGVYLAQDIKTKRMACKLIPVDQFKPSDV 187
QY 442 MACAGLTSPIRVLYGAVREGPWNIEMELLEGGSLGQLVKEOGCLPEDRALYYLQGALE 501

Db 188 EIQAQCFRHENIAELYGAVLWGETVHLFMEAGEGSGVLEKLESCGPMREFEIIVWTKHVLK 247
QY 502 GLEYLHRSRILHGDVADNVLLSSDGSAAALCDFGHAVCLQDPDGLKSLLTGDYIP---- 557
Db 248 GLDFLHSHKVIHHDIKPSNIVFMS--TKAVLVDVFGLSVQMTED-----VYFPKDLR 296
QY 558 GTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF---RGPLCLKIA-SEPP 613
Db 297 GTEIYMSPEVILCRGHSTKADIYSLGATLIHMOTGTPPWVKRYPRSAVPSYLYIIHKQAP 356
QY 614 PVREIIPSCAPLTAQAOIEGLRKEPIHRVSAALGGKVNRAALQQVGGGLKSPWRGEYKEPR 673
Db 357 PLEDIADDCSPGMRLEIEASLERPNHPRADL-----LK----- 392
QY 674 HPPPNQANYHOTLHAQPRELSPPRAGPRAEETTGRAKLPQPLPPEPPENKSPPLTUS 733
Db 393 -----HEALN-PPREDQPRC---QSLDSALLERKLLSRKELELPENIADSSCTGS 439
QY 734 KEESGM 739
Db 440 TEESEM 445
RESULT 10
US-09-423-890-2
; Sequence 2, Application US/09423890
; Patent No. 6312934
; GENERAL INFORMATION:
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEKK PROTEIN AND NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: CPI-085CPPC
; CURRENT APPLICATION NUMBER: US/09/423,890
; CURRENT FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: USN 60/078,153
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: USN 60/099,165
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-890-2
Query Match 7.1%; Score 358.5; DB 3; Length 1302;
Best Local Similarity 23.8%; Pred. No. 3.1e-18;
Matches 158; Conservative 82; Mismatches 276; Indels 149; Gaps 24;
QY 77 AISIIAQAECECENQEFSPFSEIFIAQSKQYQSESLDQIPNNVAHATEGKMARVCKWG 136
Db 688 SLQALAPASCLSENSLEHTVHREKTGKLSATRLSASSESDISRLAGVSLGPSSTTTEQ 747
QY 137 KR---RSKARKKRRKSKSLAHAGVALAKPLRTPQECSCTIPVQEDSEPLGAPYVNT 193
Db 748 PKPAVQTKGRPHSQCLNSSPLSHA--QLMPAPAPAPCSSAFSPV---DIS-----KHR 795
QY 194 PQTFLKPLKEPGLQGLCFKQGLGEGRLPALPRSELHKLISPLQ--CLNHVWKLHHPDGGPL 251
Db 796 PQAFVPCIKPS-----ASPTQ-RKFSLQFORNCSEH-----RDSQL 832
QY 252 PLPHTPPYSPLPHPPFPPIQPMKPHLESFLKCLACVDSQKLPDPHLSKLA----- 305
Db 833 ---SPVFTQSRPPSSNIHRPKSRPVP-----GSTSKLGATKSMTLDLGS 878
QY 306 ---CVDS-----PKPLPGPH-----LEPSC-LSRGAHEKFSVEEVLVHALQGSVSS-- 347
Db 879 SRCDDSFSGGSGNNAVIPSDETFTVEDKRLDVNTELNSSIEDLLEASMPSSDTTIVT 938
QY 348 -----SOAHSLSLAKTWAAGRSR---EPSP 372
Db 939 FKSEVAVLSEKAENDDTYKDDVNNHNOCKEKEEAEALAIAMAMWSQDALPIVPL 998

Db 203 ---SPVFTQSRPPSSNIHRPKPSRPV-----GSTSKLGDKATKSMMLDLGSA 248
Qy 306 ---CVDS-----PKPLPGPH-----LEPSC-LSRGAHEKFSVEEYLVHALQGSVSS-- 347
Db 249 SRCDDSFSGGNGSNAVIPSDETVPTVEDKCRLDVNTLNSIEDLLEASMPSSDITVT 308
Qy 348 -----SOAHSLSLAKTWAARGSR--EPSP 372
Db 309 FKSEVAVLSPEKAENDTYKDDVNNHKNQCKEKEEAEALAIAMWASQDALPIV 368
Qy 373 KTEDEGVLLEKLP-----VDYREBEVHWATHQLRGRSGFGEVHRMEDKOTGF 424
Db 369 QVENGEDIIIIQDTPETLPHTKAKQPYREDAEWLKGQ-QIGLGFSSCYQAQDVGTGT 427
Qy 425 QCAVKV-----RLEVFA--BELMACAGLTSPIVPLYGAVREGPWVNIFFMELLE 473
Db 428 LMAVKQVTVYRNTSSQEVEEVEALREIRMMGHLNHPNIIRMLGATCEKSNYLFIEWMA 487
Qy 474 GGSGLQVLKQCGCLPEDRALYILGQALEGLYLSRRILHGDVADNVLLSSDGSAAALC 533
Db 488 GGSVAHLLSKYGAFKESVVIINTEQLLRGLSYLHENQIHRDVKGANLLIDSTGQRLIA 547
Qy 534 DFHGAVALCLQDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGC 593
Db 548 DFGAARLASKGTGAGEFQGLL-GTIAFMAPEVLRGQYGRSCDVWVGCAIEMACAK 606
Qy 594 HPWTQFFRG---PLCLKIASE-----PPVREIPPPCAPLTAQ--AIQGLRKEP 638
Db 607 PPWNAEKHSNHLALIFKIASATTAPSIHSLSPGLRDVAVRCLQLQPDQPPSRELLKHP 666
Qy 639 IHRVS 643
Db 667 VFRTT 671

RESULT 14

US-08-461-146C-2
; Sequence 2, Application US/08461146C
; Patent No. 5981265
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHODS FOR REGULATING MEKK PROTEIN ACTIVITY
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive and Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,146C
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/354,516
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/049,254
; FILING DATE: 15-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/323,460
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/11690
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04178

RESULT 15

US-08-461-145C-2
; Sequence 2, Application US/08461145C
; Patent No. 6074861

; FILING DATE: 15-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: KARA, Catherine J.
; REGISTRATION NUMBER: P41,106
; REFERENCE/DOCKET NUMBER: CPI-004CN3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 672 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-146C-2

Query Match 7.0%; Score 355.5; DB 2; Length 672;
Best Local Similarity 23.6%; Pred. No. 2.1e-18;
Matches 157; Conservative 81; Mismatches 278; Indels 149; Gaps 24;

Qy 77 AISIIAQAECENSQEFSPFTSERIFIAGSKQYSQSESLDQIPNNVAHATEGKMARVCWK 136
Db 58 SLQAVATPSCLENSSLHTVHREKTGKLSATRLSASSEDISDRLAGVSVGLPSSTTTEQ 117
Qy 137 KR---RSKARKRKSKSKSLAHAGVALAKPLPRTPEQESCTIPVQEDESPILGAPYVNT 193
Db 118 PKPAVQTKGRPHSQCLNSPLSHA--QLMFPAPAPCSPSPV---DIS-----KHR 165
Qy 194 POFTKPLKEGCLGQLCFKLGEGRLPALPRSELHKLISPLQ--CLNVHVKLHHHDDGGPL 251
Db 166 POAFVPCIPS-----ASPQY--RKFSLQFORNCSEH-----RDSQL 202
Qy 252 PLTPHPFPYRLPHFPFPHPLQPMKPHLESFLKACVDSQKPLPDPHLSKLA----- 305
Db 203 ---SPVFTQSRPPSSNIHRPKPSRPV-----GSTSKLGDKATKSMMLDLGSA 248
Qy 306 ---CVDS-----PKPLPGPH-----LEPSC-LSRGAHEKFSVEEYLVHALQGSVSS-- 347
Db 249 SRCDDSFSGGNGSNAVIPSDETVPTVEDKCRLDVNTLNSIEDLLEASMPSSDITVT 308
Qy 348 -----SOAHSLSLAKTWAARGSR--EPSP 372
Db 309 FKSEVAVLSPEKAENDTYKDDVNNHKNQCKEKEEAEALAIAMWASQDALPIV 368
Qy 373 KTEDEGVLLEKLP-----VDYREBEVHWATHQLRGRSGFGEVHRMEDKOTGF 424
Db 369 QVENGEDIIIIQDTPETLPHTKAKQPYREDAEWLKGQ-QIGLGFSSCYQAQDVGTGT 427
Qy 425 QCAVKV-----RLEVFA--BELMACAGLTSPIVPLYGAVREGPWVNIFFMELLE 473
Db 428 LMAVKQVTVYRNTSSQEVEEVEALREIRMMGHLNHPNIIRMLGATCEKSNYLFIEWMA 487
Qy 474 GGSGLQVLKQCGCLPEDRALYILGQALEGLYLSRRILHGDVADNVLLSSDGSAAALC 533
Db 488 GGSVAHLLSKYGAFKESVVIINTEQLLRGLSYLHENQIHRDVKGANLLIDSTGQRLIA 547
Qy 534 DFHGAVALCLQDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGC 593
Db 548 DFGAARLASKGTGAGEFQGLL-GTIAFMAPEVLRGQYGRSCDVWVGCAIEMACAK 606
Qy 594 HPWTQFFRG---PLCLKIASE-----PPVREIPPPCAPLTAQ--AIQGLRKEP 638
Db 607 PPWNAEKHSNHLALIFKIASATTAPSIHSLSPGLRDVAVRCLQLQPDQPPSRELLKHP 666
Qy 639 IHRVS 643
Db 667 VFRTT 671

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 12:02:31 ; Search time 119.819 Seconds
(without alignments)
2846.907 Million cell updates/sec

Title: US-09-155-676B-7

Perfect score: 5052

Sequence: 1 MAVMEMACPGAPGSAVCGQK.....PDGSFAWSRVKHCQLENRP 947

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599520 seqs, 360203123 residues

Total number of hits satisfying chosen parameters: 1599520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5052	100.0	947	9 US-09-871-889-1	Sequence 1, Appli
2	5052	100.0	947	10 US-09-981-397A-18	Sequence 18, Appl
3	5052	100.0	947	15 US-10-394-322A-44	Sequence 44, Appl
4	5040	99.8	947	13 US-10-087-192-888	Sequence 888, App
5	2481.5	49.1	548	13 US-10-087-192-888	Sequence 885, App
6	369	7.3	891	15 US-10-425-114-57758	Sequence 57758, A
7	367	7.3	250	10 US-09-976-782-31	Sequence 31, Appl
8	367	7.3	250	15 US-10-042-865-102	Sequence 102, App
9	367	7.3	251	15 US-10-052-648A-41	Sequence 41, Appl
10	367	7.3	252	15 US-10-074-978A-160	Sequence 160, App
11	367	7.3	254	15 US-10-074-978A-150	Sequence 150, App
12	367	7.3	255	15 US-10-087-684-98	Sequence 98, Appl

13	367	7.3	255	15	US-10-218-779-98	Sequence 98, Appl
14	367	7.3	256	11	US-09-964-956-72	Sequence 72, Appl
15	256	7.3	256	15	US-10-055-569A-99	Sequence 99, Appl
16	367	7.3	256	15	US-10-042-865-83	Sequence 83, Appl
17	367	7.3	256	15	US-10-072-012-799	Sequence 799, App
18	367	7.3	256	15	US-10-072-012-856	Sequence 856, App
19	367	7.3	256	15	US-10-072-012-876	Sequence 876, App
20	367	7.3	256	15	US-10-072-012-882	Sequence 882, App
21	367	7.3	256	15	US-10-029-020-183	Sequence 183, App
22	360	7.1	467	15	US-10-287-226-24	Sequence 24, Appl
23	360	7.1	467	15	US-10-287-226-36	Sequence 36, Appl
24	360	7.1	468	15	US-10-287-226-8	Sequence 8, Appli
25	360	7.1	471	15	US-10-287-226-18	Sequence 18, Appl
26	360	7.1	474	15	US-10-287-226-4	Sequence 4, Appli
27	360	7.1	474	15	US-10-287-226-26	Sequence 26, Appl
28	360	7.1	475	15	US-10-287-226-16	Sequence 16, Appl
29	360	7.1	475	15	US-10-287-226-32	Sequence 32, Appl
30	360	7.1	478	15	US-10-287-226-10	Sequence 10, Appl
31	360	7.1	486	15	US-10-287-226-20	Sequence 20, Appl
32	360	7.1	486	15	US-10-287-226-22	Sequence 22, Appl
33	359	7.1	351	16	US-10-664-421-155	Sequence 155, App
34	359	7.1	467	9	US-09-374-579-4	Sequence 4, Appli
35	359	7.1	467	15	US-10-374-414-4	Sequence 4, Appli
36	359	7.1	467	15	US-10-262-511-228	Sequence 228, App
37	359	7.1	467	15	US-10-287-226-2	Sequence 2, Appli
38	358.5	7.1	1302	13	US-10-000-864-2	Sequence 2, Appli
39	358.5	7.1	1511	15	US-10-433-794-6	Sequence 6, Appli
40	358.5	7.1	1511	16	US-10-618-941-109	Sequence 109, App
41	358	7.1	256	10	US-09-863-776-40	Sequence 40, Appl
42	357	7.1	1512	15	US-10-210-130-64	Sequence 64, Appl
43	356.5	7.1	423	15	US-10-287-226-30	Sequence 30, Appl
44	356.5	7.1	424	15	US-10-287-226-34	Sequence 34, Appl
45	356.5	7.1	425	15	US-10-287-226-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1

US-09-871-889-1
; Sequence 1, Application US/09871889
; Patent No. US20020042499A1
; GENERAL INFORMATION:
; APPLICANT: Greene, Warner C.
; APPLICANT: Lin, Xin
; APPLICANT: Gelezuinas, Romas
; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED
; FILE REFERENCE: 30448.61USUI
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US/09/871.889
; PRIOR FILING DATE: 1999-02-25
; PRIOR APPLICATION NUMBER: 60/076,299
; PRIOR FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 947
; TYPE: PRT
; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)
US-09-871-889-1

Query Match 100.0%; Score 5052; DB 9; Length 947;
Best Local Similarity 100.0%; Pred. No. 1.8e-286; Indels 0; Gaps 0;
Matches 947; Conservative 0; Mismatches 0;
Qy 1 MAVMEMACPGAPGSAVCGQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEI 60
Db 1 MAVMEMACPGAPGSAVCGQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEI 60
Qy 61 VITKGTAKESGAPPAISIAQAECNSQEFSTFSEIPFIAGSKQYSQESLDQIPNN 120


```
Db 61 VITKTAKEGSEAGPAAISIIAAECENSQEFSTFSEIRIFTAGSKQYSQSESIDQIPNN 120
QY 121 VAHATGKMARVCWKGRSKARKKSKSLAHAGVALAKPLRTPTEQESCTTPVOE 180
Db 121 VAHATGKMARVCWKGRSKARKKSKSLAHAGVALAKPLRTPTEQESCTTPVOE 180
QY 181 DESPLGAPVVRNTPOFTKPLKEPGLGQLCFKQLEGLRPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPVVRNTPOFTKPLKEPGLGQLCFKQLEGLRPALPRSELHKLISPLQCLNHVW 240
QY 241 KLHPDQGGPLPLPHTPPYSRLPHFPPLPQWPKHPLSPGLKACVDQSKPLDPDH 300
Db 241 KLHPDQGGPLPLPHTPPYSRLPHFPPLPQWPKHPLSPGLKACVDQSKPLDPDH 300
QY 301 LSKLACVDSPKPLPGPHLPSCLSRGAHEKFSVEEYLVHALQSGSVSSQAHSITSLAKTW 360
Db 301 LSKLACVDSPKPLPGPHLPSCLSRGAHEKFSVEEYLVHALQSGSVSSQAHSITSLAKTW 360
QY 361 AAGRSRSPSKPTEDNEGVLITKLPVDYREEVHWAHQRLGRGSFGEVHRMEDK 420
Db 361 AAGRSRSPSKPTEDNEGVLITKLPVDYREEVHWAHQRLGRGSFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVPLYGAVREGPWVNI FMELEGGSLGOL 480
Db 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVPLYGAVREGPWVNI FMELEGGSLGOL 480
QY 481 VKEQGCLPEDRALYILGOALEGLYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGCLPEDRALYILGOALEGLYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Db 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
QY 601 RGPLCLKIASEPPPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAAEELGCKVNRALQQVGG 660
Db 601 RGPLCLKIASEPPPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAAEELGCKVNRALQQVGG 660
QY 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPRPAEETTGRAPKLQPLPPE 720
Db 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPRPAEETTGRAPKLQPLPPE 720
QY 721 PPEPNKSPPLTLSKEESGMWPEPLSSLEPAPARNPSSPERKATVPQEQLQLELEFLN 780
Db 721 PPEPNKSPPLTLSKEESGMWPEPLSSLEPAPARNPSSPERKATVPQEQLQLELEFLN 780
QY 781 SLSQPSLEEQEQLISCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSWSQAARSS 840
Db 781 SLSQPSLEEQEQLISCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSWSQAARSS 840
QY 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGIISSQIPAAAFS 900
Db 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGIISSQIPAAAFS 900
QY 901 LVTQKQPVRYDMEVPDSDGIDLQCTLAPDGSFAWSNRVKHGLENR 947
Db 901 LVTQKQPVRYDMEVPDSDGIDLQCTLAPDGSFAWSNRVKHGLENR 947

RESULT 2
US-09-981-397A-18
; Sequence 18, Application US/09981397A
; Publication No. US20030082519A1
; GENERAL INFORMATION:
; APPLICANT: Axixima Pharmaceuticals AG
; APPLICANT: Schubart, Daniel
; APPLICANT: Habenberger, Peter
; APPLICANT: Stein-Gerlach, Matthias
; APPLICANT: Bevec, Dorian
; TITLE OF INVENTION: Cellular Kinases Involved in Cytomegalovirus Infection and their
; TITLE OF INVENTION: Inhibition
; FILE REFERENCE: AXM-004.1 US
; CURRENT APPLICATION NUMBER: US/09/981,397A
```

```
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/240,750
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-981-397A-18

Query Match 100.0%; Score 5052; DB 10; Length 947;
Best Local Similarity 100.0%; Pred. No. 1.8e-286;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVMEMACPGAGSGAVGQOKELPKPKKKTTPPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAGSGAVGQOKELPKPKKKTTPPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKTAKEGSEAGPAAISIIAAECENSQEFSTFSEIRIFTAGSKQYSQSESIDQIPNN 120
Db 61 VITKTAKEGSEAGPAAISIIAAECENSQEFSTFSEIRIFTAGSKQYSQSESIDQIPNN 120
QY 121 VAHATGKMARVCWKGRSKARKKSKSLAHAGVALAKPLRTPTEQESCTTPVOE 180
Db 121 VAHATGKMARVCWKGRSKARKKSKSLAHAGVALAKPLRTPTEQESCTTPVOE 180
QY 181 DESPLGAPVVRNTPOFTKPLKEPGLGQLCFKQLEGLRPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPVVRNTPOFTKPLKEPGLGQLCFKQLEGLRPALPRSELHKLISPLQCLNHVW 240
QY 241 KLHPDQGGPLPLPHTPPYSRLPHFPPLPQWPKHPLSPGLKACVDQSKPLDPDH 300
Db 241 KLHPDQGGPLPLPHTPPYSRLPHFPPLPQWPKHPLSPGLKACVDQSKPLDPDH 300
QY 301 LSKLACVDSPKPLPGPHLPSCLSRGAHEKFSVEEYLVHALQSGSVSSQAHSITSLAKTW 360
Db 301 LSKLACVDSPKPLPGPHLPSCLSRGAHEKFSVEEYLVHALQSGSVSSQAHSITSLAKTW 360
QY 361 AAGRSRSPSKPTEDNEGVLITKLPVDYREEVHWAHQRLGRGSFGEVHRMEDK 420
Db 361 AAGRSRSPSKPTEDNEGVLITKLPVDYREEVHWAHQRLGRGSFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVPLYGAVREGPWVNI FMELEGGSLGOL 480
Db 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVPLYGAVREGPWVNI FMELEGGSLGOL 480
QY 481 VKEQGCLPEDRALYILGOALEGLYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGCLPEDRALYILGOALEGLYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Db 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
QY 601 RGPLCLKIASEPPPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAAEELGCKVNRALQQVGG 660
Db 601 RGPLCLKIASEPPPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAAEELGCKVNRALQQVGG 660
QY 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPRPAEETTGRAPKLQPLPPE 720
Db 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPRPAEETTGRAPKLQPLPPE 720
QY 721 PPEPNKSPPLTLSKEESGMWPEPLSSLEPAPARNPSSPERKATVPQEQLQLELEFLN 780
Db 721 PPEPNKSPPLTLSKEESGMWPEPLSSLEPAPARNPSSPERKATVPQEQLQLELEFLN 780
QY 781 SLSQPSLEEQEQLISCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSWSQAARSS 840
Db 781 SLSQPSLEEQEQLISCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSWSQAARSS 840
QY 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGIISSQIPAAAFS 900
Db 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGIISSQIPAAAFS 900
```

```
Db      841  SNNVLAGRPTDTSYFNGVKVQIQSLNGEHLHIREFHRVVKVGDIAIGSIQIPAAAFS 900
Qy      901  LVTKDQPVRYDMEVPSGIDLQCTLAPDGSFAMSRVVKHGOLENRP 947
Db      901  LVTKDQPVRYDMEVPSGIDLQCTLAPDGSFAMSRVVKHGOLENRP 947

RESULT 3
US-10-394-322A-44
; Sequence 44, Application US/10394322A
; Publication No. US20030232391A1
; GENERAL INFORMATION:
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.
; APPLICANT: Prescott, John C.
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS
; FILE REFERENCE: 39750-0006 US
; CURRENT APPLICATION NUMBER: US/10/394,322A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/366,892
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-394-322A-44

Query Match      100.0%; Score 5052; DB 15; Length 947;
Best Local Similarity 100.0%; Pred. No. 1.8e-286;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAVMEMACPGAGSAGVQKQKELPKPKXTPLPGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db      1  MAVMEMACPGAGSAGVQKQKELPKPKXTPLPGKQSSVYKLEAVEKSPVFCGKWEILND 60

Qy      61  VITKGTAKGSEAGPAALIIIAQACENSQSFPTFSRIFIAQSKQYSESLEQIPNN 120
Db      61  VITKGTAKGSEAGPAALIIIAQACENSQSFPTFSRIFIAQSKQYSESLEQIPNN 120

Qy      121  VAHATEGMARVCKWKGKRSKARKKRSKSLAHAGVALAKPLPRTPEQESCTIPVOE 180
Db      121  VAHATEGMARVCKWKGKRSKARKKRSKSLAHAGVALAKPLPRTPEQESCTIPVOE 180

Qy      181  DESPLGAPYVRNTPOFTKPLKEPGLQGLGEGLRPALPRSELHKLISPLQCLNHVW 240
Db      181  DESPLGAPYVRNTPOFTKPLKEPGLQGLGEGLRPALPRSELHKLISPLQCLNHVW 240

Qy      241  KLHPDQGGPLPLPHTPPYSLRPLPFPPLQPKWPHLESFLGKLACVDSQKPLDPH 300
Db      241  KLHPDQGGPLPLPHTPPYSLRPLPFPPLQPKWPHLESFLGKLACVDSQKPLDPH 300

Qy      301  LSKLACVDSQKPLDPHLEPCLSRGAHEKFSVEYLVAHQSGSVSSQAHSLSLAKTW 360
Db      301  LSKLACVDSQKPLDPHLEPCLSRGAHEKFSVEYLVAHQSGSVSSQAHSLSLAKTW 360

Qy      361  AARGSRSEPKTDENGVLLTEKLKPDVYEEVEVHWATHQLRGLGRSGFGEVHRMEDK 420
Db      361  AARGSRSEPKTDENGVLLTEKLKPDVYEEVEVHWATHQLRGLGRSGFGEVHRMEDK 420

Qy      421  QTGFQCAVKVRLEVFRAEELMACAGLTPSRIVPLYGAVREGPWVNIIFMELLEGSLGQL 480
Db      421  QTGFQCAVKVRLEVFRAEELMACAGLTPSRIVPLYGAVREGPWVNIIFMELLEGSLGQL 480

Qy      481  VKEOGLPEDRALYYLGOALEGLYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db      481  VKEOGLPEDRALYYLGOALEGLYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540

Qy      541  LQPDGLGSLTLDGYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600
Db      541  LQPDGLGSLTLDGYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600

Qy      601  RGPLCLKIASPPVPREIPPSCAPLTAQAIQEGLRKEPIHRVSAELGKVNRLAQOVGG 660
```

```
Db      601  RGPLCLKIASPPVPREIPPSCAPLTAQAIQEGLRKEPIHRVSAELGKVNRLAQOVGG 660
Qy      661  LKSPWRGEYKEPRHPPPNQANYHOTLHAQRELSPRAPGPABETTGAPKIQPLPPE 720
Db      661  LKSPWRGEYKEPRHPPPNQANYHOTLHAQRELSPRAPGPABETTGAPKIQPLPPE 720

Qy      721  PPEPNKSPPLTLSKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQLEIFLN 780
Db      721  PPEPNKSPPLTLSKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQLEIFLN 780

Qy      781  SLSQPFSLSEOEQILSCLSIDSLSDSEKXNPASKOSSRDTLSSGVHSSQAEARSS 840
Db      781  SLSQPFSLSEOEQILSCLSIDSLSDSEKXNPASKOSSRDTLSSGVHSSQAEARSS 840

Qy      841  SNNVLAGRPTDTSYFNGVKVQIQSLNGEHLHIREFHRVVKVGDIAIGSIQIPAAAFS 900
Db      841  SNNVLAGRPTDTSYFNGVKVQIQSLNGEHLHIREFHRVVKVGDIAIGSIQIPAAAFS 900

Qy      901  LVTKDQPVRYDMEVPSGIDLQCTLAPDGSFAMSRVVKHGOLENRP 947
Db      901  LVTKDQPVRYDMEVPSGIDLQCTLAPDGSFAMSRVVKHGOLENRP 947

RESULT 4
US-10-087-192-888
; Sequence 888, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 888
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-888

Query Match      99.8%; Score 5040; DB 13; Length 947;
Best Local Similarity 99.8%; Pred. No. 9.2e-286;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  MAVMEMACPGAGSAGVQKQKELPKPKXTPLPGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db      1  MAVMEMACPGAGSAGVQKQKELPKPKXTPLPGKQSSVYKLEAVEKSPVFCGKWEILND 60

Qy      61  VITKGTAKGSEAGPAALIIIAQACENSQSFPTFSRIFIAQSKQYSESLEQIPNN 120
Db      61  VITKGTAKGSEAGPAALIIIAQACENSQSFPTFSRIFIAQSKQYSESLEQIPNN 120

Qy      121  VAHATEGMARVCKWKGKRSKARKKRSKSLAHAGVALAKPLPRTPEQESCTIPVOE 180
Db      121  VAHATEGMARVCKWKGKRSKARKKRSKSLAHAGVALAKPLPRTPEQESCTIPVOE 180

Qy      181  DESPLGAPYVRNTPOFTKPLKEPGLQGLGEGLRPALPRSELHKLISPLQCLNHVW 240
Db      181  DESPLGAPYVRNTPOFTKPLKEPGLQGLGEGLRPALPRSELHKLISPLQCLNHVW 240

Qy      241  KLHPDQGGPLPLPHTPPYSLRPLPFPPLQPKWPHLESFLGKLACVDSQKPLDPH 300
Db      241  KLHPDQGGPLPLPHTPPYSLRPLPFPPLQPKWPHLESFLGKLACVDSQKPLDPH 300

Qy      301  LSKLACVDSQKPLDPHLEPCLSRGAHEKFSVEYLVAHQSGSVSSQAHSLSLAKTW 360
```

Db 301 LSKLACVDSKPLPGPHLESPCLSRGAHEKFSVEEYLVHALQSVSSGQAHSLTSLAKTW 360
Qy 361 AARGSRSPKTEDENEGVLTEKLPVDYEVREEVHATHQLRLGRSGFGEVHMEK 420
Db 361 AARGSRSPKTEDENEGVLTEKLPVDYEVREEVHATHQLRLGRSGFGEVHMEK 420
Qy 421 QTGFQCAVKVRLEVFRAELMACAGLTSPRIVPLYGAVREGPWVNI FMELEGGSLGQL 480
Db 421 QTGFQCAVKVRLEVFRAELMACAGLTSPRIVPLYGAVREGPWVNI FMELEGGSLGQL 480
Qy 481 VKQGLCPEDRALYYLGOALEGLEYLHRSRILHGDVKADNVLLSDGSHAAALCDFGHAYC 540
Db 481 VKQGLCPEDRALYYLGOALEGLEYLHRSRILHGDVKADNVLLSDGSHAAALCDFGHAYC 540
Qy 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMHLMLNGCHPWTOFF 600
Db 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMHLMLNGCHPWTOFF 600
Qy 601 RGPLCLKIASPPVREIIPSCAPLTAQAIQEGLRKEPIHRVSAABELGKGNRALQOVGG 660
Db 601 RGPLCLKIASPPVREIIPSCAPLTAQAIQEGLRKEPIHRVSAABELGKGNRALQOVGG 660
Qy 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPPRAGPRPAEETTGRAKLPPLPPE 720
Db 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPPRAGPRPAEETTGRAKLPPLPPE 720
Qy 721 PPEPNKSPPLTUSKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQOLELEFLN 780
Db 721 PPEPNKSPPLTUSKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQOLELEFLN 780
Qy 781 SLSPQFSLSEEQILSCLSIDSLSDSEKPNKSPKASQSSRDTLSGSHWSQAERSS 840
Db 781 SLSPQFSLSEEQILSCLSIDSLSDSEKPNKSPKASQSSRDTLSGSHWSQAERSS 840
Qy 841 SNWVLARGRPDTSPYFNGVKVQIQLNGEHLHIREFHVRKVGDIATGISSQIPAAAFS 900
Db 841 SNWVLARGRPDTSPYFNGVKVQIQLNGEHLHIREFHVRKVGDIATGISSQIPAAAFS 900
Qy 901 LVTKGQPVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKVGQLENRP 947
Db 901 LVTKGQPVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKVGQLENRP 947

RESULT 5

US-10-087-192-885
; Sequence 885, Application US/10087192
; Publication No. US20020182386A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 885
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-885

Query Match 49.1%; Score 2481.5; DB 13; Length 548;
Best Local Similarity 86.1%; Pred. No. 1.2e-136;
Matches 466; Conservative 32; Mismatches 36; Indels 7; Gaps 2;
Qy 407 GRGSFGEVHRMEDKDTGFQCAVKVRLEVFRAELMACAGLTSPRIVPLYGAVREGPWVN 466

Db 15 GQSSFGVHRMKDKDTGFQCAVKVRLEVFRAELMACAGLTSPRIVPLYGAVREGPWVN 74
Qy 467 IFMELLEGSLGQLVKBOGCLPEDRALYYLGOALEGLEYLHRSRILHGDVKADNVLLSSD 526
Db 75 IFMELLEGSLGQLVKBOGCLPEDRALYYLGOALEGLEYLHRSRILHGDVKADNVLLSSD 134
Qy 527 GSHAALCDFGHAYCLOPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCM 586
Db 135 GSHAALCDFGHAYCLOPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCM 194
Qy 587 LHMINGCHPWTOFFRGPLCLKIASPPVREIIPSCAPLTAQAIQEGLRKEPIHRVSAAE 646
Db 195 LHMINGCHPWTOFFRGPLCLKIASPPVREIIPSCAPLTAQAIQEGLRKEPIHRVSAAE 254
Qy 647 LGKGNRALQOVGGUKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPPRAGPRPAEET 706
Db 255 LRRKVGKALQEVGGLKSPWKGEYKEPRPPQDQATCHOTLPTPRE-NP-----PAKAN 307
Qy 707 TGRAPKLOPPLPPEPPEPNKSPPLTUSKEESGMWEPPLSSLEPAPARNPSSPERKATVP 766
Db 308 TDGAPPEPPLPPEPPEPNKSPALMLSKESGTWEPLPLSSLDLPATAKGFPPDRATLP 367
Qy 767 EQELQOLELEFLNLSLQSFSLSEEQILSCLSIDSLSDSEKPNKSPKASQSSRDTLS 826
Db 368 EQELQOLELEFLNLSLQSFSLSEEQILSCLSIDSLSDSEKPNKSPKASQSSRDTLS 427
Qy 827 GVHSSWSQAERSSWNWVLARGRPDTSPYFNGVKVQIQLNGEHLHIREFHVRKVGDI 886
Db 428 GVHSSWSQAERSSWNWVLARGRPDTSPYFNGVKVQIQLNGEHLHIREFHVRKVGDI 487
Qy 887 ATGSSQIPAAAFSLVTKDGOPVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKVGQLENR 946
Db 488 ATGSSQIPAAAFSLVTKDGOPVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKVGQLENR 547
Qy 947 P 947
Db 548 P 548

RESULT 6

US-10-425-114-57758
; Sequence 57758, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 57758
; LENGTH: 891
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GMFLMINSOY034B02_FLI.fli
US-10-425-114-57758

Query Match 7.3%; Score 369; DB 15; Length 891;
Best Local Similarity 23.4%; Pred. No. 4.5e-13;
Matches 224; Conservative 107; Mismatches 353; Indels 272; Gaps 43;
Qy 88 NSQSFSPFSEIRIFIA-----GSKQYSQ---SESIDQIPNNVAHATEGKMARVCW 134
Db 21 NKESFFDTLHRKLRISKGVKSVKSGSRHRCNTISEKGDHSPCGSRSPSKVAR-CQ 79

D8: 229 KLLVKDPEKRLTAE 244

```
RESULT 10
US-10-074-978A-160
; Sequence 160, Application US/10074978A
; Publication No. US20040010119A1
; GENERAL INFORMATION:
; APPLICANT: Leite, Mario
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Fernandes, Elma
; APPLICANT: Li, Li
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Liu, Xiahong
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Patturajan, Meera
; APPLICANT: Blalock, Angela
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Gusev, Vladimir
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter S
; APPLICANT: Ellerman, Karen
; APPLICANT: Heyes, Melvin P
; APPLICANT: Herrman, John
; APPLICANT: Pena, Carol E A
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Moore, No. US20040010119Alle
; APPLICANT: Shenoy, Suresh
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, Dave
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glenda
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Moore, No. US20040010119Alle
; APPLICANT: Shenoy, Suresh
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, Dave
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glenda
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-269
; CURRENT APPLICATION NUMBER: US/10/074,978A
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: 60/268,221
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/335,109
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/312,284
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/268,496
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/276,703
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/330,293
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/322,127
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/280,899
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/310,797
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/268,646
; PRIOR FILING DATE: 2001-02-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 160
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: S TKC,
```

OTHER INFORMATION: Serine/Threonine protein kinases domain sequence
US-10-074-978A-160

```
Query Match 7.3%; Score 367; DB 15; Length 252;
Best Local Similarity 37.9%; Pred. No. 1.3e-13;
Matches 97; Conservative 37; Mismatches 94; Indels 28; Gaps 8;

QY 406 LGRGSFGVHRMEDKOTGFQCACVKKVRLEVFRAB-----ELMACAGLTSPRIVPLVGA 458
|:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||
DB 7 LGKGAFKVYLARDKTKTGKLVAIKVIKBEKILKKKRERILREITKILKLDHPNIVKLYDV 66
|:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

QY 459 VREGPWNIIFMELLEGLGSLGQLVKEOCLPEDRALYVLGOALSGLEVYLHSRRILHGDVKA 518
|:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||
DB 67 FEDDDKLYLVMEYCEGGDLFDLLKKRGJSEDFARFYAQILSALRYLHSGIIRDLKP 126
|:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

QY 519 DNVLLSSDGSHAAALCDFGHAVCIQPDGLGKSLTLTDGYICTETHMAPEVVVLRSCDAKVD 578
|:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||
DB 127 ENILLDSOG-HVKLADFLGAKQLDSGG---TLILT--TFVGTPTPYMAPEVLLGKYGVKAND 180
|:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

QY 579 VWSSCCMLHLMNGCHPWTQTFFRG-----PLCLKIASBPPVPVREIPPP---SCAPLTAQAI 630
|:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||
DB 181 IWSLGVLTYELLTGKPP-----FPGDQQLLALFKKIGKPPPPP---FPPPEWKISPEAKDLI 233
|:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

QY 631 QEGLRKEPIHRVSAAE 646
|:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||
DB 234 KOLLVKDPEKRLTAE 249
|:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||
```

RESULT 11
US-10-074-978A-150
; Sequence 150, Application US/10074978A
; Publication No. US20040010119A1
; GENERAL INFORMATION:
; APPLICANT: Leite, Mario
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Fernandes, Elma
; APPLICANT: Li, Li
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Liu, Xiahong
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Patturajan, Meera
; APPLICANT: Blalock, Angela
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Gusev, Vladimir
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter S
; APPLICANT: Ellerman, Karen
; APPLICANT: Heyes, Melvin P
; APPLICANT: Herrman, John
; APPLICANT: Pena, Carol E A
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Moore, No. US20040010119Alle
; APPLICANT: Shenoy, Suresh
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, Dave
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glenda
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-269
; CURRENT APPLICATION NUMBER: US/10/074,978A
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: 60/268,221
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/335,109
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/312,284
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/268,496
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/276,703
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/330,293
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/322,127
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/280,899
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/310,797
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/268,646
; PRIOR FILING DATE: 2001-02-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 160
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: S TKC,

```
; PRIOR APPLICATION NUMBER: 60/312,284
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/268,496
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/276,703
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/330,293
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/322,127
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/280,899
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/310,797
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/268,646
; PRIOR FILING DATE: 2001-02-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 150
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: s_ktc,
; OTHER INFORMATION: Serine/Threonine protein kinases
US-10-074-978A-150

Query Match          7.3%; Score 367; DB 15; Length 254;
Best Local Similarity 37.9%; Pred. No. 1.3e-13;
Matches 97; Conservative 37; Mismatches 94; Indels 28; Gaps 8;

QY 406 LGRSGFGEVHRMEDKQTGFCQAVKVRLEVFRAE-----ELMACAGLTSPRIVLYGA 458
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 7 LGKGAFGKYVLARDKTKGLVAIKVKIKKXKREILRIKILKLDHPNIVKLYDV 66

QY 459 VREGPWNIIFMELLEGSLGQLVKEQGLPEDRALYVLCQALEGLYLSRRLHGDVKA 518
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 67 FEDDDKLYLVMEYCEGGDFDLKKRGRLSEDEARFYARQILSALEYLHSGQIIHRDLKP 126

QY 519 DNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDIYIPGTETHMAPEVILGRSCDAKD 578
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 127 ENILLSDSG-HVKLADFGAKQLDSGG---TLT--TFVGTPEYMAPEVILGKGYGKAVD 180

QY 579 VNSSCCMLHMLNGCHPWTFQFRG-----PLCLKIASPPVREIPP---SCAPLTAQAI 630
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 181 IWSGLVILYELLTKGPP---FPGDDQLLALFKKIGKPPPP---FPPPEWKISPEAKDLI 233

QY 631 QEGLRKEPIHRVSAE 646
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 234 KLLVKDPEKRLTAEE 249

RESULT 12
US-10-087-684-98
; Sequence 98, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catharine E.
; APPLICANT: Casman, Stacie L.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
```

```
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles E.
; APPLICANT: Gangolli, Esha A.
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-214 CIP
; CURRENT APPLICATION NUMBER: US/10/087,684
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 98
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Domain
US-10-087-684-98

Query Match          7.3%; Score 367; DB 15; Length 255;
Best Local Similarity 37.9%; Pred. No. 1.4e-13;
Matches 97; Conservative 37; Mismatches 94; Indels 28; Gaps 8;

QY 406 LGRSGFGEVHRMEDKQTGFCQAVKVRLEVFRAE-----ELMACAGLTSPRIVLYGA 458
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 6 LGKGAFGKYVLARDKTKGLVAIKVKIKKXKREILRIKILKLDHPNIVKLYDV 65

QY 459 VREGPWNIIFMELLEGSLGQLVKEQGLPEDRALYVLCQALEGLYLSRRLHGDVKA 518
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 66 FEDDDKLYLVMEYCEGGDFDLKKRGRLSEDEARFYARQILSALEYLHSGQIIHRDLKP 125

QY 519 DNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDIYIPGTETHMAPEVILGRSCDAKD 578
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 126 ENILLSDSG-HVKLADFGAKQLDSGG---TLT--TFVGTPEYMAPEVILGKGYGKAVD 179

QY 579 VNSSCCMLHMLNGCHPWTFQFRG-----PLCLKIASPPVREIPP---SCAPLTAQAI 630
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 180 IWSGLVILYELLTKGPP---FPGDDQLLALFKKIGKPPPP---FPPPEWKISPEAKDLI 232

QY 631 QEGLRKEPIHRVSAE 646
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 233 KLLVKDPEKRLTAEE 248

RESULT 13
US-10-218-779-98
; Sequence 98, Application US/10218779
; Publication No. US2004002922A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit
; APPLICANT: MacDougall, John
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William
; APPLICANT: Alsbrook II, John
```


QY 579 VWSCCMMLHMLNGCHPWTQFFRG-----PLCLKIASPPPPVREIPP---SCAPLTAQAI 630
Db 181 IWSLGVILYELLTGKPP-----FPGDDQLLALFKKIGKPPPP---FPPPEWKISPEAKDLI 233

QY 631 QEGLRKEPIHRVSAAE 646
Db 234 KKLVLKDPKRLTAEE 249

RESULT 15
US-10-055-569A-99
; Sequence 99, Application US/10055569A
; Publication No. US20040024181A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli, Esha A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gilbert, Jennifer
; APPLICANT: Casman, Stacie
; APPLICANT: Bialock, Angela
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Shenoy, Suresh
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit
; APPLICANT: Malyanker, Uriel
; APPLICANT: Stone, David
; APPLICANT: Millet, Isabelle
; APPLICANT: Smithson, Glenna
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Taupier Jr., Raymond J
; APPLICANT: Anderson, David W
; TITLE OF INVENTION: No. US20040024181A1el Human Proteins, Polynucleotides Encoding Th
; FILE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 21402-191
; CURRENT APPLICATION NUMBER: US/10/055,569A
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 60/243,642
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,320
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,592
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,681
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/243,863
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/244,443
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 60/245,029
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,995
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/245,293
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/245,315
; PRIOR FILING DATE: 2000-11-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 99
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-055-569A-99

Query Match 7.3%; Score 367; DB 15; Length 256;
Best Local Similarity 37.9%; Pred. No. 1.4e-13;
Matches 97; Conservative 37; Mismatches 94; Indels 28; Gaps 8;

QY 406 LGRGSFGEVHRMEDKQTFQCAVKKVRLEVFRAE-----ELMACAGLTSPRIVPLYGA 458
Db 7 LGKGAFGKVIYARDKKTGKLVAKVIKKEKLKKKRERILREIKILKKLDHPNIVKLYDV 66

QY 459 VREGFWYNIIFMELLEGGSLGOLVKEQCLPEDRALYVLGOALEGLELYLHSSRIILHGDVKA 518
Db 67 FEDDDKLYLVMEYCEGGDLFDLLKKRGLSEDEARFYARQILSALEYLHSGQIIRDLKP 126

QY 519 DNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDIYIPGTETHMAPEVVLGRSCDAKVD 578
Db 127 ENILLSDG-HVKLADFGLAKQLDSGG---TLLT--TFVGTPEYMAPEVVLGKGYGKAVD 180

QY 579 VWSCCMMLHMLNGCHPWTQFFRG-----PLCLKIASPPPPVREIPP---SCAPLTAQAI 630
Db 181 IWSLGVILYELLTGKPP-----FPGDDQLLALFKKIGKPPPP---FPPPEWKISPEAKDLI 233

QY 631 QEGLRKEPIHRVSAAE 646
Db 234 KKLVLKDPKRLTAEE 249

Search completed: June 7, 2005, 12:25:35
Job time : 121.819 secs

Result No.	Score	Query Match	Length	DB	ID	Description	
1	1133.4	43.1	3156	2	US-08-887-518-1	Sequence 1, Appli	
2	1133.4	43.1	3156	2	US-09-023-321-1	Sequence 1, Appli	
3	1133.4	43.1	3156	2	US-09-032-475-1	Sequence 1, Appli	
4	977	37.1	2844	3	US-09-257-703-2	Sequence 2, Appli	
5	977	37.1	2844	4	US-09-871-889A-2	Sequence 2, Appli	
6	42.4	1.6	825	4	US-09-252-991A-9023	Sequence 9023, Ap	
7	42.4	1.6	1359	4	US-09-252-991A-9108	Sequence 9108, Ap	
C	8	42.4	1.6	1746	4	US-09-252-991A-9885	Sequence 9885, Ap
	9	40.2	1.5	835	6	5223425-7	Patent No. 5223425
10	40.2	1.5	835	6	5223425-7	Patent No. 5223425	
11	40.2	1.5	1093	6	5223425-3	Patent No. 5223425	
12	40.2	1.5	1093	6	5223425-3	Patent No. 5223425	
13	40	1.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli	
14	40	1.5	4411529	3	US-09-103-840A-1	Sequence 1, Appli	
15	39.6	1.5	1005	1	US-08-287-442-8	Sequence 8, Appli	
16	39.6	1.5	1005	1	US-08-459-701-8	Sequence 8, Appli	
17	39.6	1.5	1005	1	US-08-460-298-8	Sequence 8, Appli	
18	39.6	1.5	1005	1	US-08-459-174-8	Sequence 8, Appli	
19	39.2	1.5	430	4	US-09-621-976-16566	Sequence 16566, A	
C	20	39	1.5	7218	1	US-08-232-463-14	Sequence 14, Appl
	21	38.8	1.5	1233	4	US-09-949-016-3456	Sequence 3456, Ap
22	38.8	1.5	1612	1	US-08-343-733A-1	Sequence 1, Appli	
23	38.4	1.5	399	4	US-09-621-976-8976	Sequence 8976, Ap	
24	38.4	1.5	408	4	US-09-252-991A-6166	Sequence 6166, Ap	
25	38.4	1.5	470	4	US-10-101-464A-422	Sequence 422, App	
26	38.4	1.5	1371	4	US-09-252-991A-6010	Sequence 6010, Ap	
C	27	38.4	1.5	1710	4	US-09-252-991A-6299	Sequence 6299, Ap

Db	1928	CTCAGCGAGCTGGGAGGAGGTGAACCGGCACTACAGCAAGTGGAGGTCTGAAGA	1987
Qy	121	GCCTTTGAGGGGAGAAATATAAGAAACAAAGACATCCACCGCCAAATCAAGCCAAATTACC	180
Db	1988	GCCTTTGAGGGGAGAAATATAAGAAACAAAGACATCCACCGCCAAATCAAGCCAAATTACC	2047
Qy	181	ACCAGACCTTCATCCCGAGCGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGGGCCAG	240
Db	2048	ACCAGACCTTCATCCCGAGCGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGGGCCAG	2107
Qy	241	CTGAGGAGACAAACAGGAGAGCCCTTAAGCTTCAGCTTCCTCTCCACAGAGCCCGCAG	300
Db	2108	CTGAGGAGACAAACAGGAGAGCCCTTAAGCTTCAGCTTCCTCTCCACAGAGCCCGCAG	2167
Qy	301	AGCCAAAACAAGTCTCCTCCCTTGACTTTGAGCAAGGAGAGTCTGGGATGTGGGAACCT	360
Db	2168	AGCCAAAACAAGTCTCCTCCCTTGACTTTGAGCAAGGAGAGTCTGGGATGTGGGAACCT	2227
Qy	361	TACCTCTGTCTCCTCGAGCGAGCCCTCGCCAGAAACCCAGCTTCACAGAGCGGAAAG	420
Db	2228	TACCTCTGTCTCCTCGAGCGAGCCCTCGCCAGAAACCCAGCTTCACAGAGCGGAAAG	2287
Qy	421	CAACCGTCCCGAGCAGGAGTGCAGCAGCTGGAATAGAAATTTCTCTCAACAGCCTGT	480
Db	2288	CAACCGTCCCGAGCAGGAGTGCAGCAGCTGGAATAGAAATTTCTCTCAACAGCCTGT	2347
Qy	481	CCAGCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTCCTCAGCATCGACGCC	540
Db	2348	CCAGCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTCCTCAGCATCGACGCC	2407
Qy	541	TCTCCCTGTGCGATGACATGAGAGAAACCCATCAAGGCCCTCTCAAGCTTCGCGGACA	600
Db	2408	TCTCCCTGTGCGATGACATGAGAGAAACCCATCAAGGCCCTCTCAAGCTTCGCGGACA	2467
Qy	601	CCCTGAGCTCAGCGTACACTCTCTGAGCAGCAGCGCCGAGCTCGAAGCTCCAGCTGGA	660
Db	2468	CCCTGAGCTCAGCGTACACTCTCTGAGCAGCAGCGCCGAGCTCGAAGCTCCAGCTGGA	2527
Qy	661	ACATGTGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGTGTGAAG	720
Db	2528	ACATGTGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGTGTGAAG	2587
Qy	721	TCCAAATACAGTCTCTTAATGTGAACACCTGACATCCGGAGTTCACCGGGTCAAG	780
Db	2588	TCCAAATACAGTCTCTTAATGTGAACACCTGACATCCGGAGTTCACCGGGTCAAG	2647
Qy	781	TGGGAGACATCGGCATCGGCATCAGCAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA	840
Db	2648	TGGGAGACATCGGCATCGGCATCAGCAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA	2707
Qy	841	CCAAGACGGGAGCCTTTCGCTACGACATGAGGTGCAGACTCGGGGATCGACCTGC	900
Db	2708	CCAAGACGGGAGCCTTTCGCTACGACATGAGGTGCAGACTCGGGGATCGACCTGC	2767
Qy	901	AGTGACACTGGCCCTGATGGCAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC	960
Db	2768	AGTGACACTGGCCCTGATGGCAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC	2827
Qy	961	TGGAGAACAGGCCTTAACCTTCGCTCCACCGCGGCTCCACACTCGCGGAAAGCAGCCT	1020
Db	2828	TGGAGAACAGGCCTTAACCTTCGCTCCACCGCGGCTCCACACTCGCGGAAAGCAGCCT	2887
Qy	1021	TCCTGTCTGGTGAACAGATGTGCTGCTGAAACACAGCTCAGCGTTCAGGGGATYTG	1080
Db	2888	TCCTGTCTGGTGAACAGATGTGCTGCTGAAACACAGCTCAGCGTTCAGGGGATYTG	2945
Qy	1081	NCCAGCCCCCGGCTCAGCAGTGGGAACAGGGCTTCGNCAGCNAGCNAGTNGGGG	1140
Db	2946	GCCAGCCCCCGGCTCAGCAGTGGGAACAGGGCTTCGNCAGCNAGCNAGTNGGGG	2998
Qy	1141	CAAGCNAGATTCCTCCAGGATTTCAACNCTGAGCCNCTGCCNANCCCTCTCTGAADA	1200
Db	1201	AAACATVTCGCCACCTGTAAGAGACAGAGGAGTGCNCAGGATTNACCTYGGGGAA	1260
Qy	3053	AAAAACATCCGCCACCTGTAAGAGACAGAGGAGGATGGC---AGGAGTTACCTGGGAA	3108
Db	1261	ACAAAACAGGAGTCTTTTCTGCTCCCTGTCTCCAGTNCAGTTCGGCTG	1309
Qy	3109	ACAAAACAGGAGTCTTTTCTGCTCCCTGTCTCCAGTTCGGCTG	3155
Db	1868	CCCCCTCTCACAGCCCGAGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGT	1927
Qy	61	CTGACGCGAGCTGGGAGGAGGTGAACCCGGGCACTACAGCAAGTGGGAGTCTGAAGA	120
Db	1928	CTGACGCGAGCTGGGAGGAGGTGAACCCGGGCACTACAGCAAGTGGGAGTCTGAAGA	1987
Qy	121	GCCTTTGAGGGGAGAAATATAAGAAACAAAGACATCCACCGCCAAATCAAGCCAAATTACC	180
Db	1988	GCCTTTGAGGGGAGAAATATAAGAAACAAAGACATCCACCGCCAAATCAAGCCAAATTACC	2047
Qy	181	ACCAGACCTTCATCCCGAGCGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGGGCCAG	240

RESULT 2

US-09-023-321-1
; Sequence 1, Application US/09023321
; Patent No. 5844073
; GENERAL INFORMATION:
; APPLICANT: Rothe, Mike
; APPLICANT: Wu, Lin
; TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,321
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/887,518
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3156 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-09-023-321-1

Query Match 43.1%; Score 1133.4; DB 2; Length 3156;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 1241; Conservative 4; Mismatches 43; Indels 21; Gaps 7;

Db 2048 ACAGACCTCCATGCCAGCGGAGAGAGCTTTCCGCAAGGGCCCAAGGGCCCGGCCAG 2107
Qy 241 CTGAGGAGACAAAGGAGAGCCCTTAAGCTTCAAGCCTCTCTCCACCAAGAGCCGCCAG 300
Db 2108 CTGAGGAGACAAAGGAGAGCCCTTAAGCTTCAAGCCTCTCTCCACCAAGAGCCGCCAG 2167
Qy 301 AGCCAAACAAGTCTCTCCCTTGACTTTGACAAAGGAGGAGTCTGGGATGTGGAAACCT 360
Db 2168 AGCCAAACAAGTCTCTCCCTTGACTTTGACAAAGGAGGAGTCTGGGATGTGGAAACCT 2227
Qy 361 TACCTCTGTCTCTCCCTCGAGCAGCCCTCTGCAAAACCCAGCTCACAGAGCGGAAAG 420
Db 2228 TACCTCTGTCTCTCCCTCGAGCAGCCCTCTGCAAAACCCAGCTCACAGAGCGGAAAG 2287
Qy 421 CAACCGTCCCGAGCAGGAATCGCAGCAGCTGGAATAGAAATATTCTCTCAACAGCCTGT 480
Db 2288 CAACCGTCCCGAGCAGGAATCGCAGCAGCTGGAATAGAAATATTCTCTCAACAGCCTGT 2347
Qy 481 CCCAGCAATTTCTCTGGAGAGCAGGAGCAAAATCTCTGTGCTCAGCATCGACAGCC 540
Db 2348 CCCAGCAATTTCTCTGGAGAGCAGGAGCAAAATCTCTGTGCTCAGCATCGACAGCC 2407
Qy 541 TCTCCCTGTGGATGACAGTGAAGAACCCATCAAGAGCCTCTCAAGCTCGCGGACA 600
Db 2408 TCTCCCTGTGGATGACAGTGAAGAACCCATCAAGAGCCTCTCAAGCTCGCGGACA 2467
Qy 601 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGCTCGAAGCTCCAGCTGGA 660
Db 2468 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGCTCGAAGCTCCAGCTGGA 2527
Qy 661 ACATGGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 720
Db 2528 ACATGGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 2587
Qy 721 TCCAAATACAGTCTCTTAATGGTGAAACCTCTGCAATCCGGGAGTTTCCACGGGTCAAAG 780
Db 2588 TCCAAATACAGTCTCTTAATGGTGAAACCTCTGCAATCCGGGAGTTTCCACGGGTCAAAG 2647
Qy 781 TGGGAGACATCGCACTGGCATCAGCAGCAGATCCAGCTGCGAGCTTCAAGCTTGTGCA 840
Db 2648 TGGGAGACATCGCACTGGCATCAGCAGCAGATCCAGCTGCGAGCTTCAAGCTTGTGCA 2707
Qy 841 CCAAAGACGGCAGCCTGTTGCTACGACATGAGAGTGCAGACTCGGGCATCGACCTGC 900
Db 2708 CCAAAGACGGCAGCCTGTTGCTACGACATGAGAGTGCAGACTCGGGCATCGACCTGC 2767
Qy 901 AGTGCACATCGGCCCTGATGGAGCTTGGCTCGAGCTGAGGCTCAAGTGGCCAGC 960
Db 2768 AGTGCACATCGGCCCTGATGGAGCTTGGCTCGAGCTGAGGCTCAAGTGGCCAGC 2827
Qy 961 TGGAGACAGGCCCTTAACCTCGCCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT 1020
Db 2828 TGGAGACAGGCCCTTAACCTCGCCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT 2887
Qy 1021 TCCTGCTCGGTGACGATGCTGCTGAAACACAGAGCTCAGCCGTTTCCAGGGGATYTG 1080
Db 2888 TCCTGCTCGGTGACGATGCTGCTGAAACACAGAGCTCAGCCGTTTCCAGGGGATY-- 2945
Qy 1081 NCCAGCCCCCGGTTCARCAAGTGGGAACAGGCGCTCGNAGCAGNAGNAGTNGGGG 1140
Db 2946 GCCAGCCCCCGGTTC--CAGTGGGAACAGGCGCTCG-----CAGCAGCAAGGTGGG 2998
Qy 1141 CAAGNAGAAATGCTCCAGATTTTCAACNCTGAGCCCTGCCCCCAGCCCTGCTGAADA 1200
Db 2999 GCAAGCAGAAATGCTCCAGATTTTCAACCTGAGCCCTGCCCCCAGCCCTGCTGA 3052
Qy 1201 AAAAAYTNCGCCCACTGAAGAGACAGAAAGGAGATGNCAGGATTTNACCTTGGGAA 1260
Db 3053 AAAAATATCCGCACGCTGAAGAGACAGAAAGGAGATGNC-----AGGATTTACCTGGGAA 3108
Qy 1261 ACAAAACAGGATCTTTTTCCTGCCCCCTGCTCCAGTNCAGTGTGGCCTG 1309
Db 3109 ACAAAACAGGATCTTTT--TTCTGCCCCCTGCTCCAGT--CGAGTTGGCCTG 3155

RESULT 3

US-09-032-475-1
; Sequence 1, Application US/09032475
; Patent No. 5854003
; GENERAL INFORMATION:
; APPLICANT: Roche, Mike
; APPLICANT: Wu, Lin
; TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,475
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/887,518
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3156 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-032-475-1

Query Match 43.1%; Score 1133.4; DB 2; Length 3156;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 1241; Conservative 4; Mismatches 43; Indels 21; Gaps 7;

Qy 1 CCCCTCTCACAGCCCGAGCCATCCAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT 60
Db 1868 CCCCTCTCACAGCCCGAGCCATCCAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT 1927
Qy 61 CTCGAGCGGAGCTGGGAGGGAAGGTGAACCGGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120
Db 1928 CTCGAGCGGAGCTGGGAGGGAAGGTGAACCGGGGCACTACAGCAAGTGGGAGGTCTGAAGA 1987
Qy 121 GCCCTTGGAGGGGAGATATAAAGAACCAAGACATCCACCGCCCAATCAAGCCAAATTACC 180
Db 1988 GCCCTTGGAGGGGAGATATAAAGAACCAAGACATCCACCGCCCAATCAAGCCAAATTACC 2047
Qy 181 ACCAGACCCCTCCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGGCCAG 240
Db 2048 ACCAGACCCCTCCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGGCCAG 2107
Qy 241 CTGAGGAGACAAAGGAGGAGCCCTTAAGCTTCCAGCCTCTCTCCCAAGAGAGCCGCCAG 300
Db 2108 CTGAGGAGACAAAGGAGGAGCCCTTAAGCTTCCAGCCTCTCTCTCCCAAGAGAGCCGCCAG 2167
Qy 301 AGCCAAACAAGTCTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGGAAACCT 360
Db 2168 AGCCAAACAAGTCTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGGAAACCT 2227

```
QY 361 TACCTCTGTCTCTCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACCAGACGGGAAG 420
Db 2228 TACCTCTGTCTCTCTCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACCAGACGGGAAG 2287
QY 421 CAACCGTCCGGAGCAGGAACCTGCAGCAGCTGGAATAGAAATATCTCTCAACAGCTGT 480
Db 2288 CAACCGTCCGGAGCAGGAACCTGCAGCAGCTGGAATAGAAATATCTCTCAACAGCTGT 2347
QY 481 CCAGCATTCTCTCTGGAGGAGCAGGAGCAATCTCTCGTGCCCTCAGCATCGACAGCC 540
Db 2348 CCAGCATTCTCTCTGGAGGAGCAGGAGCAATCTCTCGTGCCCTCAGCATCGACAGCC 2407
QY 541 TCTCCCTGTCCGATGACAGTGAAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGACA 600
Db 2408 TCTCCCTGTCCGATGACAGTGAAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGACA 2467
QY 601 CCTGAGCTCAGCGGTACACTCTCTGGAGAGCCAGCCAGAGCTCGAAGCTCCAGCTGA 660
Db 2468 CCTGAGCTCAGCGGTACACTCTCTGGAGAGCCAGCCAGAGCTCGAAGCTCCAGCTGA 2527
QY 661 ACATGCTGTGCGCCGGGGGGGGCCCAACGACACCCCAAGCTATTTCAATGGTGAAG 720
Db 2528 ACATGCTGTGCGCCGGGGGGGGCCCAACGACACCCCAAGCTATTTCAATGGTGAAG 2587
QY 721 TCCAAATACAGTCTCTTAATGGTGAACACTCTGCACATCCGGGAGTTCCACCGGGTCAAAG 780
Db 2588 TCCAAATACAGTCTCTTAATGGTGAACACTCTGCACATCCGGGAGTTCCACCGGGTCAAAG 2647
QY 781 TGGAGACATCGGCATCGGCATCAGAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA 840
Db 2648 TGGAGACATCGGCATCGGCATCAGAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA 2707
QY 841 CAAAGAGCGGCGAGCTGTTCCTAGACATGAGGTGCAGACTCGGGCATCGACCTGC 900
Db 2708 CAAAGAGCGGCGAGCTGTTCCTAGACATGAGGTGCAGACTCGGGCATCGACCTGC 2767
QY 901 AGTGACACTGGCCCTCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGCGCCAGC 960
Db 2768 AGTGACACTGGCCCTCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGCGCCAGC 2827
QY 961 TGGAGAACAGGCGCTTAACCTGCCCTCCACCGCGGCTCCACACTGCGGGAAGCAGCT 1020
Db 2828 TGGAGAACAGGCGCTTAACCTGCCCTCCACCGCGGCTCCACACTGCGGGAAGCAGCT 2887
QY 1021 TCTGTCTGCTCAGATGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATYG 1080
Db 2888 TCTGTCTGCTCAGATGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATYG 2945
QY 1081 NCCAGCCCCCGGCTCARCAGNTGGGAACAGGGCTCTGNCAGCNAGNAGGTTNGGGGG 1140
Db 2946 GCCAGCCCCCGGCTCA--CAGTGGGAACAGGGCTCG----CAGCAGCAAGGTGGGG 2998
QY 1141 CAAGCNAGATGCTCCAGGATTTTCANCTGAGCCNTGCCCAACCTGCTGAADA 1200
Db 2999 GCAAGCAGATGCTCCAGGATTTTCACACTGAGCCCTGCCCCAC-----CCTGTGTA 3052
QY 1201 AAACATVTCGCGCAGTGAAGAGACAGAGGAGGATGNCAGGAGTTNACCTYGGGAA 1260
Db 3053 AAACATVTCGCGCAGTGAAGAGACAGAGGAGGATGNCAGGAGTTNACCTYGGGAA 3108
QY 1261 ACAAACAGGATCTTTTCTGCCCCCTGCTCCAGTNCAGGTTGGCCTG 1309
Db 3109 ACAAACAGGATCTTTTCTGCCCCCTGCTCCAGT-CGAGTTGGCCTG 3155
```

RESULT 4

US-09-257-703-2

; Sequence 2, Application US/09257703

; Patent No. 626538

; GENERAL INFORMATION:

; APPLICANT: Greene, Warner C.

; APPLICANT: Lin, Xin

```
; APPLICANT: Gelezuinas, Romas
; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED
; FILE REFERENCE: 30448.61USU1
; CURRENT APPLICATION NUMBER: US/09/257,703
; EARLIER APPLICATION NUMBER: 60/076,299
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; TYPE: DNA
; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)
US-09-257-703-2
```

```
Query Match 37.1%; Score 977; DB 3; Length 2844;
Best Local Similarity 100.0%; Pred. No. 6.3e-272;
Matches 977; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCTCTCACAGCCAGGCCCATCCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
Db 1868 CCCCTCTCACAGCCAGGCCCATCCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 1927
QY 61 CTGACAGCGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTTGAAGA 120
Db 1928 CTGACAGCGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTTGAAGA 1987
QY 121 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 180
Db 1988 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 2047
QY 181 ACCAGACCTCCATGCCCAGCCGAGAGAGCTTTCGCAAGGGGCCCGGGCCCGCAG 240
Db 2048 ACCAGACCTCCATGCCCAGCCGAGAGAGCTTTCGCAAGGGGCCCGGGCCCGCAG 2107
QY 241 CTGAGGAGACAAACAGGAGAGCCCTTAAGCTCCAGCCCTCTCTCCACAGAGCCCCAG 300
Db 2108 CTGAGGAGACAAACAGGAGAGCCCTTAAGCTCCAGCCCTCTCTCCACAGAGCCCCAG 2167
QY 301 AGCCAAACAAGTCTCTCCCTTGAATTTGACAGAGGAGGTCTGGGATGTGGGAACCT 360
Db 2168 AGCCAAACAAGTCTCTCCCTTGAATTTGACAGAGGAGGTCTGGGATGTGGGAACCT 2227
QY 361 TACCTCTGCTCTCCTCGAGCAGCCCTGCGCAGAAACCCAGCTCACCAGAGCGGAAG 420
Db 2228 TACCTCTGCTCTCCTCGAGCAGCCCTGCGCAGAAACCCAGCTCACCAGAGCGGAAG 2287
QY 421 CAACCGTCCCGAGCAGGAACCTGACAGAGCTGGAAATAGAAATATTCTCTCAACAGCTGT 480
Db 2288 CAACCGTCCCGAGCAGGAACCTGACAGAGCTGGAAATAGAAATATTCTCTCAACAGCTGT 2347
QY 481 CCGAGCCATTTTCTCTGGAGGAGCAGGAGCAAAATCTCTCGTGCTCTCAGCATCGACGCC 540
Db 2348 CCGAGCCATTTTCTCTGGAGGAGCAGGAGCAAAATCTCTCGTGCTCTCAGCATCGACGCC 2407
QY 541 TCTCCCTGTCCGATGACAGTGAAGAAACCCATCAAGGGCTCTCAAGCTCGCGGACA 600
Db 2408 TCTCCCTGTCCGATGACAGTGAAGAAACCCATCAAGGGCTCTCAAGCTCGCGGACA 2467
QY 601 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCAGGCGAGGCTCGAAGCTCCAGCTGA 660
Db 2468 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCAGGCGAGGCTCGAAGCTCCAGCTGA 2527
QY 661 ACATGGTGTGGCCCGGGGGGGCCCAACCCGACACCCCAAGCTATTTCAATGGTGAAG 720
Db 2528 ACATGGTGTGGCCCGGGGGGGCCCAACCCGACACCCCAAGCTATTTCAATGGTGAAG 2587
QY 721 TCCAAATACAGTCTCTTAATGGTGAACACTCTGACATCCGGGAGTTCCACCGGGTCAAAG 780
Db 2588 TCCAAATACAGTCTCTTAATGGTGAACACTCTGACATCCGGGAGTTCCACCGGGTCAAAG 2647
QY 781 TGGAGACATCGGCATCGGCATCAGAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA 840
```


Db 719 GCGGGCAGAGCGCGCGGAGCCCTTCATCGACCTCAACGACAGCTACCCCGTGGTGCAG 778

QY 864 TAGCAGATCGAGGTGCCAGACTCGGGCATCGACCTG 899

Db 779 GCGGACATGATCTCGGTGCCTCGGGCTCAAGCCG 814

RESULT 7

US-09-252-991A-9108

; Sequence 9108, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 9108

; LENGTH: 1359

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-9108

Query Match 1.6%; Score 42.4; DB 4; Length 1359;

Best Local Similarity 54.5%; Pred. No. 0.13;

Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 744 GAACACCTGCACATCGGGAGTTCCACCGGTCAAAGTGGGAGACATCGCCACTGGCATC 803

Db 610 GAACACATGGCCAAAGCAGCCCTGCGAGCAGTGGCAAGCCGATGAAGATCGCCATTATCGGC 669

QY 804 AGCAGCGATGCCAGCTGCAGCTTCCAGCTTGGTCAACAAAGACGGGCGAGCCTGTTCGC 863

Db 670 GCGGGCAGAGCGCGCGGAGCCCTTCATCGACCTCAACGACAGCTACCCGTGGTGCAG 729

QY 864 TAGCAGATCGAGGTGCCAGACTCGGGCATCGACCTG 899

Db 730 GCGGACATGATCTCGGTGCCTCGCGCTCAAGCCG 765

RESULT 8

US-09-252-991A-8985/c

; Sequence 8985, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 8985

; LENGTH: 1746

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-8985

Query Match 1.6%; Score 42.4; DB 4; Length 1746;

Best Local Similarity 54.5%; Pred. No. 0.15;

Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 744 GAACACCTGCACATCGGGAGTTCCACCGGTCAAAGTGGGAGACATCGCCACTGGCATC 803

Db 438 GAACACATGGCCAAGCAGCCCTGCAGCAGTGGCAAGCCGATGAAGATGCCATTATCGGC 379

QY 804 AGCAGCCAGATCCAGCTGCAGCCTTCACTTGTGTCCACAAAGACGGGCGACCTGTTGCG 863

Db 378 GCGGGCAGAGCGCGGAGGCTTTCATCGACCTCAACGACAGCTACCCGTGGTGCAG 319

QY 864 TAGCAGATCGAGGTGCCAGACTCGGGCATCGACCTG 899

Db 318 GCGGACATGATCTCGGTGCCTCGGGCTCAAGCCG 283

RESULT 9

5223425-7

; Patent No. 5223425

; APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,

; BARRY M.; WHITE, TYLER R.

; TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT

; D ACTIVITY

; NUMBER OF SEQUENCES: 19

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/277,963

; FILING DATE: 30-NOV-1988

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 34,203

; FILING DATE: 02-APR-1987

; SEQ ID NO: 7

; LENGTH: 835

5223425-7

Query Match 1.5%; Score 40.2; DB 6; Length 835;

Best Local Similarity 48.1%; Pred. No. 0.43;

Matches 114; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 747 CACCTGCACATCCGGGAGTTCCACCGGTCAAAGTGGGAGACATCGCCACTGGCATCAGC 806

Db 494 GCGCGGACAGCTGCGACGCTCTTGCAGTGTGGACCGCGCACCTGCAACCGG 553

QY 807 AGCCAGATCCAGCTGCAGCCTTCAGCTTGTCCACAAAGACGGGCGAGCCTGTTGCTAC 866

Db 554 CGCAGCACACGACGCGGCCATCAGAGCGCTTGATGCGCGGAGAGCAATCGCCGG 613

QY 867 GACATGGAGTGCAGACTCGGGCATCGACCTGCAGTGCACACTGGCCCTTGATGGCAGC 926

Db 614 GACAGCTCAAGGCTGACTCCGGGGCGCGCTGTGTGCGGGGCGTGCTCGAGGGCGTG 673

QY 927 TTCGCTGGAGCTCGAGGTCAAGCATGGCAGCTGGAGAACAGCCCTTAACCTGCG 983

Db 674 GTCACCTCGGCTCGCGGTTTTCGGGCAACCGCAAGACCGCGGATCTACACCGCG 730

RESULT 10

5223425-7

; Patent No. 5223425

; APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,

; BARRY M.; WHITE, TYLER R.

; TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT

; D ACTIVITY

; NUMBER OF SEQUENCES: 19

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/277,963

; FILING DATE: 30-NOV-1988

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 34,203

; FILING DATE: 02-APR-1987

; SEQ ID NO: 7

; LENGTH: 835

5223425-7

Query Match

Best Local Similarity 48.1%; Pred. No. 0.43; Length 835;

Matches 114; Conservative 0; Mismatches 123; Indels 0; Gaps 0;


```

; PROR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,284
; FILING DATE: 02-JUL-1992
; PROR APPLICATION DATA:
; APPLICATION NUMBER: US 07/570,184
; FILING DATE: 08-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1005 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; STRAIN: single
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1005
; OTHER INFORMATION: /product= "alternate sigma factor
; OTHER INFORMATION: RpoS"
;
; US-08-287-442-8
;
Query Match 1.5%; Score 39.6; DB 1; Length 1005;
Best Local Similarity 46.5%; Pred.No. 0.72; Indels 0; Gaps 0;
Matches 126; Conservative 0; Mismatches 145;

QY 783 GGAGACATCGCCCACTGGCATCAGACGCCAGATCCCAGCTGCAGCCTTCAGCTTGTGTCACC 842
DB 601 GNAGAAATCGCCAACTGCTGGAGAAACCGGTAGGTGAGGTCAAGCGCATGCTGGGFTCTC 660
QY 843 AAGACGGGAGCGCTGTTTCGTACGACATGGAGGTGCCAGACTCGGCATCGACCTGCAG 902
DB 661 AATGAGCGGGTGTTCTTCAGTGTCAGCTTCTCGCTGGTCCGATTCGGNATAAACCCCTGCTG 720
QY 903 TGCACACTGCGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGFTCAAGCATTGGCCAGCTG 962
DB 721 GACACCTCACCGACGATCGCCAAACCGATCCGTCGAGCTGCTGCAGGATGACGACCTG 780
QY 963 GAGAACAGGCGCTTAACCTGTGCGCTCCACCGCGGCTCCACACTGCCCGAAGCAGCCTTC 1022
DB 781 TCGCAAAAGCATCGATCATGTGCTTTCGAACTGACCGACGAGTGAAGTAGTGT 840
QY 1023 CTGCTCGGTGCACGATGCTGCCCTGAAAAACA 1053
DB 841 CGCCGCTTCGGCTTCGGCGCCATGAAAGCA 871

Search completed: June 10, 2005, 20:46:45
Job time : 303.303 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 14:23:14 ; Search time 999.861 Seconds
(without alignments)

16311.995 Million cell updates/sec

Title: US-09-155-676B-3

Perfect score: 2631

Sequence: 1 cccctctcacagccagcc.....accctgcacctcagggggg 2631

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications NA:*
- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
 - 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
 - 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
 - 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
 - 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
 - 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
 - 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
 - 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
 - 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
 - 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
 - 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
 - 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
 - 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
 - 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
 - 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
 - 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
 - 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
 - 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
 - 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
 - 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
 - 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
 - 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
 - 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
 - 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
 - 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
 - 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1850.2	70.3	4596	10	US-09-981-397A-17
2	1850.2	70.3	4596	17	Sequence 17, Appl
3	1850.2	70.3	4596	18	US-10-172-118-975
4	1850.2	70.3	4596	19	Sequence 975, App
5	1783.4	67.8	4684	13	US-10-342-887-975
6	977	37.1	2844	9	US-10-283-975A-234
7	973.4	37.0	73967	13	US-10-087-192-887
					Sequence 2, Appli
					Sequence 886, App

8	754.2	28.7	1385	18	US-10-296-115-97	Sequence 97, Appl
9	672.8	25.6	2828	13	US-10-087-192-884	Sequence 884, App
c 10	349.4	13.3	2475	17	US-10-108-260A-1754	Sequence 1754, Ap
11	242.2	9.2	499	10	US-09-918-995-2585	Sequence 2585, Ap
12	193	7.3	296	9	US-09-783-590-11852	Sequence 11852, A
13	188.8	7.2	42566	13	US-10-087-192-883	Sequence 883, App
c 14	88.4	3.4	455	20	US-10-357-930-22052	Sequence 22052, A
c 15	88.4	3.4	455	20	US-10-357-930-27909	Sequence 27909, A
c 16	88.4	3.4	476	20	US-10-357-930-12325	Sequence 12325, A
c 17	88.4	3.4	504	20	US-10-357-930-33496	Sequence 33496, A
c 18	88.4	3.4	504	20	US-10-357-930-42420	Sequence 42420, A
c 19	73.8	2.8	463	20	US-10-357-930-31556	Sequence 31556, Ap
c 20	52.8	2.0	65	10	US-09-908-975-25736	Sequence 25736, A
21	50	1.9	50	17	US-10-131-827-2583	Sequence 2583, Ap
22	50	1.9	50	17	US-10-131-827-3841	Sequence 3841, Ap
23	49	1.9	60	10	US-09-908-975-19057	Sequence 19057, A
24	43	1.6	447	20	US-10-425-115-153036	Sequence 153036,
25	42.6	1.6	1464	20	US-10-258-148-11	Sequence 11, Appl
c 26	42.4	1.6	867	14	US-10-123-153-20	Sequence 20, Appl
c 27	42.4	1.6	867	15	US-10-156-761-5737	Sequence 5737, Ap
c 28	42.4	1.6	867	15	US-10-146-731-20	Sequence 20, Appl
c 29	42.4	1.6	867	15	US-10-140-472-20	Sequence 20, Appl
c 30	42.4	1.6	867	15	US-10-141-761-20	Sequence 20, Appl
c 31	42.4	1.6	867	16	US-10-142-885-20	Sequence 20, Appl
c 32	42.4	1.6	867	16	US-10-158-790-20	Sequence 20, Appl
c 33	42.4	1.6	867	17	US-10-137-871-20	Sequence 20, Appl
c 34	42.4	1.6	867	17	US-10-140-923-20	Sequence 20, Appl
c 35	42.4	1.6	867	17	US-10-141-756-20	Sequence 20, Appl
c 36	42.4	1.6	867	17	US-10-141-759-20	Sequence 20, Appl
c 37	42.4	1.6	867	17	US-10-140-803-20	Sequence 20, Appl
c 38	42.4	1.6	867	17	US-10-140-864-20	Sequence 20, Appl
c 39	42.4	1.6	867	18	US-10-142-426-20	Sequence 20, Appl
c 40	42.4	1.6	9025608	15	US-10-156-761-1	Sequence 1, Appl1
c 41	40.4	1.5	1089	17	US-10-282-122A-14856	Sequence 14856, A
c 42	40.2	1.5	699	9	US-09-821-255-1	Sequence 1, Appli
c 43	40.2	1.5	762	14	US-10-202-676-3	Sequence 3, Appli
c 44	40.2	1.5	835	14	US-10-202-676-1	Sequence 1, Appli
c 45	40.2	1.5	911	13	US-10-044-090-214	Sequence 214, App

ALIGNMENTS

RESULT 1

US-09-981-397A-17
; Sequence 17, Application US/09981397A
; Publication No. US20030082519A1
; GENERAL INFORMATION:
; APPLICANT: Axxima Pharmaceuticals AG
; APPLICANT: Schubart, Daniel
; APPLICANT: Habenberger, Peter
; APPLICANT: Stein-Gezlach, Matthias
; APPLICANT: Bevec, Dorian
; TITLE OF INVENTION: Cellular Kinases Involved in Cytomegalovirus Infection and their
; FILE REFERENCE: AXM-004.1 US
; CURRENT APPLICATION NUMBER: US/09/981,397A
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/240,750
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-981-397A-17

Query Match 70.3%; Score 1850.2; DB 10; Length 4596;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;
QY 1 CCCCTCTCACAGCCAGCCATCCAAAGAGGGGTGAGGAAGAGCCATCCACCGCTGT 60

Db 2100 CCCCTCTCAGAGCCAGGCCATCCAAAGAGGGCTGAGGAAGAGCCCATCCACCGGTGT 2159
QY 61 CTGACGGAGCTGGGAGGGAAGTGAACCCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120
Db 2160 CTGACGGAGCTGGGAGGGAAGTGAACCCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 2219
QY 121 GCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAATCAAGCCAAATTACC 180
Db 2220 GCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAATCAAGCCAAATTACC 2279
QY 181 ACCAGACCCCTCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCAAGGGCCCGGCCAG 240
Db 2280 ACCAGACCCCTCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCAAGGGCCCGGCCAG 2339
QY 241 CTGAGGAGACACAGGAGAGCCCTTAAGCTCCAGCCTCTCTCCACACAGAGCCCCCAG 300
Db 2340 CTGAGGAGACACAGGAGAGCCCTTAAGCTCCAGCCTCTCTCCACACAGAGCCCCCAG 2399
QY 301 AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGAGAGTCTGGGATGTGGGAACCT 360
Db 2400 AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGAGAGTCTGGGATGTGGGAACCT 2459
QY 361 TACCTCTGTCTCTCTGGAGCAGCCCTTGCCAGAAACCCAGCTCAACAGAGCGGAAAG 420
Db 2460 TACCTCTGTCTCTCTGGAGCAGCCCTTGCCAGAAACCCAGCTCAACAGAGCGGAAAG 2519
QY 421 CAAACGTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAAATATTCTCAACAGCCTGT 480
Db 2520 CAAACGTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAAATATTCTCAACAGCCTGT 2579
QY 481 CCCAGCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCC 540
Db 2580 CCCAGCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCC 2639
QY 541 TCTCCCTGTCTGATGACAGTGAGAAACCCATCAAAAGGCTCTCAAAAGCTCGCGGACA 600
Db 2640 TCTCCCTGTCTGATGACAGTGAGAAACCCATCAAAAGGCTCTCAAAAGCTCGCGGACA 2699
QY 601 CCTGAGCTCAGGCTGACACTCTTGAGCAGCAGCCGCGAGCTCGAAGCTCCAGCTGA 660
Db 2700 CCTGAGCTCAGGCTGACACTCTTGAGCAGCAGCCGCGAGCTCGAAGCTCCAGCTGA 2759
QY 661 ACATGTGTCTGGCCCGGGGCGGCCCAACGACACCCCAAGCTATTTCATTTGGTGTGAAG 720
Db 2760 ACATGTGTCTGGCCCGGGGCGGCCCAACGACACCCCAAGCTATTTCATTTGGTGTGAAG 2819
QY 721 TCCAAATACAGTCTCTTAATGTGTAACACTTGACATCCGGAGTTCCACCGGTTCAAG 780
Db 2820 TCCAAATACAGTCTCTTAATGTGTAACACTTGACATCCGGAGTTCCACCGGTTCAAG 2879
QY 781 TGGGAGACATCGCCACTGAGCATCAGCAGCAGATCCAGCTTCAGCTTCAGCTTGGTCA 840
Db 2880 TGGGAGACATCGCCACTGAGCATCAGCAGCAGATCCAGCTTCAGCTTCAGCTTGGTCA 2939
QY 841 CCAAGACGGGAGCCTGTTCGTACGACATGAGAGTGCAGACTCGGGCATCGACCTGC 900
Db 2940 CCAAGACGGGAGCCTGTTCGTACGACATGAGAGTGCAGACTCGGGCATCGACCTGC 2999
QY 901 AGTGCACTGTGGCCCTGATGGAGCTTGCCTGGAGCTGGAGGTTCAAGCATGGCCAGC 960
Db 3000 AGTGCACTGTGGCCCTGATGGAGCTTGCCTGGAGCTGGAGGTTCAAGCATGGCCAGC 3059
QY 961 TGGAGAACAGGCCCTTAACCTGCTCCACCGCGGCTCCACACTCCCGGAAAGAGCCT 1020
Db 3060 TGGAGAACAGGCCCTTAACCTGCTCCACCGCGGCTCCACACTCCCGGAAAGAGCCT 3119
QY 1021 TCCTGTCTCGGTGACAGTGTCTGCTCAAAACAACAGGCTCAGCGCTTCACAGGGATYTG 1080
Db 3120 TCCTGTCTCGGTGACAGTGTCTGCTCAAAACAACAGGCTCAGCGCTTCACAGGGATYTG 3177
QY 1081 NCCAGCCCCCGGCTCAGAGNTGGGAACAGGGCCCTGNCAGCAGNAGAGGTTGGGG 1140

Db 3178 GCCAGCCCCCGGCTCA--CAGTGGGAACAGGGGCTCG-----CAGCAGCAAGGTGGG 3210
QY 1141 CAAGCNAAGATGCTCCAGGATTTCAACNCCTGAGCCCTGCCCCANCCCTGCTGAADA 1200
Db 3231 GCAAGCAAGATGCTCCAGGATTTCAACCTGAGCCCTGCCAC-----CCTGCTGA 3284
QY 1201 AAAAYTCCGCCACGTGAAGAGACAGAGGAGATGCGNAGAGTTNNACCTTGGGGAA 1260
Db 3285 AAAAACATCCGCCACGTGAAGAGACAGAGGAGATGGC---AGGAGTTTACCTGGGGAA 3340
QY 1261 ACATAACAGGATCTTTTCTGCCCCCTGCTCCAGTNCAGTGGGCTGNCACCGCTTGG 1320
Db 3341 ACATAACAGGATCTTTTCTGCCCCCTGCTCCAGT--CGAGTTGGCTGA--CCGCTTG 3396
QY 1321 ANTCAGTACCATTTGTCAGANCAGGGGAGAGCAGCTTCCAGCCCTGGGTCAAGAGG 1380
Db 3397 GATCAGTGACCATTTGTTGGCAGA--CAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGAGG 3455
QY 1381 GTGGGAGAGCCTTTCGGGCCCTCAACCTNCCAGGCTGCTGTGAGAGTGTCAAGTGTGA 1440
Db 3456 GTGGGAGAGCCTTTCGGGCCCTCAACCT--CCAGGCTGCTGTG--AGAGTGTCAAGTGTGA 3513
QY 1441 AGGNCCTCAANCTCAGNTTTCAGTGCAAGAACAGGTNCCAGCAGTATGCCGCCCGNTA 1500
Db 3514 AGGNCCTCAAACTC---AGTTTCAGTGCAAGAACAGGT--CAGCAGTATGCCGCCCGTAG 3569
QY 1501 GGTAAANNNGGGGCGCTTNAACCCCTTGCTNGGCTTNCACCTNGGCCAGCTCAACCC 1560
Db 3570 GTTAA---GGGGGCGCTTAAACCCCTTGCTGGCCTCA--CCTGGCCAGCTCA--CCC 3621
QY 1561 CTTTGGGTGAGGGGAAAGATGCTGACCCCTGGGAAGGCTWCCCTGGTGTAGATACAC 1620
Db 3622 CTTTGGGTGAGGGGAAAGATGCTGACCCCTGGGAAGGCT--CCTTGGTGTAGATACAC 3680
QY 1621 CACACTTTTCAAGTGTGTGCAACACAGCTCCTGAGTTGACCTCTGTTTCAGCCAAAGACC 1680
Db 3681 CACACTTTTCAAGTGTGTGCAACACAGCTCCTGAGTTGACCTCTGTTTCAGCCAAAGACC 3740
QY 1681 AAAGAAGGTGTGAAGTGAAGTGTCTCATGNTCCCAAGACATGTGCCCCCTTGTGCTG 1740
Db 3741 AAAGAAGGTGTGAAGTGAAGTGTCTCATG--CCCCAGACATGTGCCCCCTTGTGCTG 3799
QY 1741 GCTACCATCTTCCCGCAGCAGCAGGCCCCGAGCCCTTCCAGCCCGCAGCTGCCCCAG 1800
Db 3800 GCTACCATCTTCCCGCAGCAGCAGGCCCCGAGCCCTTCCAGCCCGCAGCTGCCCCAG 3859
QY 1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTTGATGCAAGGATATGCTGA 1860
Db 3860 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTTGATGCAAGGATATGCTGA 3919
QY 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTCTNNAAGNNAAGCAGCAGAGAGAGCY 1920
Db 3920 CAGGAACAGTCTGTGGATGGACATGATCAGTCTNNAAGNNAAGCAGCAGAGAGAGCY 3976
QY 1921 TCCGGGCCCCCAGNCCCCCACTNATCAGTGTNCCAGCGTNGGTTNCCCAAGNAGACA 1980
Db 3977 TCCGGGCCCCCAGNCCCCCACTNATCAGTGT---CCAGCGTGTGTTCCCAAGCAGCAGC- 4032
QY 1981 GCTNCAGNCACTCANCACCTGACACTNCACTNGCCCTGCCCCCTNGCCGANGAGGTA 2040
Db 4033 -----TCAGCATCAGACTGACACTCACCTGCCCCCTGG--CCAGAGGTA 4083
QY 2041 CCGNACGGCACTTTCANCTGTATGNAACCTCAAGACATTTTCATGGCTGCCCCCTNN 2100
Db 4084 CCG-ACGGCACTTTCGACT--CTGATGACCTCAAGACATTTTCATGGCTGCCCCCTG 4140
QY 2101 GCAGGNCAGGNCAGGNCAGTGTAGNAGCATGTAAGCAANGCCAGGAGATGG 2160
Db 4141 GGGCAGGNCAGGNCAGTGTAGTGT-----GTAGGACATAGCAAGCCAGGAGATGG 4190
QY 2161 GGTGNAAGGNCACAGTCTTGTAGTGTCCANCATGTGATGTGCTTCAAACTCTTNN 2220
Db 4191 GGTGAA--GGGACACAGTCTTGTAGTGTCCA-CATGCTGTGACT--CCTCAAACTCTT- 4245

2221	Qy	NCCAGNATTTCTCTAAGAAATAGCANCCCCCTTNGCCCATTTGCCCCAGCTTAGCCTCTTCT	2280
4246	Db	--CCAGATTTCTCTAAGAAATAGCACCCCC--TTCCCCATTTGCCCAGCTTAGCCTCTTCT	4301
2281	Qy	CCCAGGGGAGCTTANCTCAGAGACTCACGTAGCATTAATTCAGCTGTGNAATCGTCAGGGG	2340
4302	Db	CCCAGGGGAGCTA--CTCAGAGACTCACGTAGCATTTAAATCAGCTGTG--AAATCGTCAGGGG	4359
2341	Qy	TGTCGTCTAGCTCAACTCTCTGGGCAGGGGAGCGCGAGAGACTCCGTGGCGAGAGCTCAT	2400
4360	Db	TGTCGTCTAGCTCAACTCTCTGGGCAGGGGAGCGCGAGAGACTCCGTGGGAGAGCTCAT	4419
2401	Qy	TCCACATCTTGCACAGACGCTTTNGTCCAGCTGTCCACATTTAGTCAGACTGCTCCC	2460
4420	Db	TCCACATCTTGCACAGACAGCTTT--GTCCAGCTGTCCACATTTAGTCAGACTGCTCCC	4478
2461	Qy	GGGGAGAGACCCCGGCCCCCGACACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTG	2520
4479	Db	GGGGAGAGACCCCGGCCCCCGACACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTG	4538
2521	Qy	GGTTGTAGAACTCTTTGTTAAGCAATAAAGTTTGGGTGATGACAAATGTTAAAAA	2578
4539	Db	GGTTGTAGAACTCTTTGTTAAGCAATAAAGTTTGGGTGATGACAAATGTTAAAAA	4596

RESULT 2

```

US-10-172-118-975
; Sequence 975, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 975
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_003954
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-975

```

Query Match	70.3%	Score 1850.2	DB 17	Length 4596
Best Local Similarity	90.5%	Pred. No. 0		
Matches 2334	Conservative	4	Mismatches 159	Indels 81
Gaps	34			
Qy 1	CCCTCTCTCACAGCCCGCCAGCCCATCAAGAGGGGGCTGAGAAAGAGCCCATCCACCGCGTGT	60		
Db 2100	CCCTCTCTCACAGCCCGCCAGCCCATCAAGAGGGGGCTGAGAAAGAGCCCATCCACCGCGTGT	2159		
Qy 61	CTCAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTCAAGA	120		
Db 2160	CTCAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTCAAGA	2219		
Qy 121	GCCTTTGGAGGGGAGAAATATAAAGAACCAAGACATCCACCGCAAAATCAAGCCAAATTACC	180		
Db 2220	GCCTTTGGAGGGGAGAAATATAAAGAACCAAGACATCCACCGCAAAATCAAGCCAAATTACC	2279		
Qy 181	ACCAGACCCCTTCATGCCCGCAGGAGAGCTTTTCGCCAAGGGGCCCCAGGGGCCCGGGCCAG	240		

Db 2460 TACCTCTGTCTCCTCGAGCCAGCCCTCTGCCAGAAACCCAGCTCACAGAGCGGAAG 2519
Qy 421 CAACCGTCCCGAGCAGGAACTGACAGAGCTGGAAATAGAAATATTCTCTCAACAGCCTGT 480
Db 2520 CAACCGTCCCGAGCAGGAACTGACAGAGCTGGAAATAGAAATATTCTCTCAACAGCCTGT 2579
Qy 481 CCAGACCATTTTCTCTGGAGGAGCAGAGCAAAATTTCTGTGTGCTCAGATCAGACGCC 540
Db 2580 CCAGACCATTTTCTCTGGAGGAGCAGAGCAAAATTTCTGTGTGCTCAGATCAGACGCC 2639
Qy 541 TCTCCCTGTGCGATGACAGTGAAGAAACCATCAAAAGGCTCTCAAAAGCTCGCGGACA 600
Db 2640 TCTCCCTGTGCGATGACAGTGAAGAAACCATCAAAAGGCTCTCAAAAGCTCGCGGACA 2699
Qy 601 CCCTGAGCTCAGGCGGTACACTCTCTGGAGCAGCAGCGCGAGGCTCGAAGCTTCCAGCTGGA 660
Db 2700 CCCTGAGCTCAGGCGGTACACTCTCTGGAGCAGCAGCGCGAGGCTCGAAGCTTCCAGCTGGA 2759
Qy 661 ACATGCTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGCTGTAAG 720
Db 2760 ACATGCTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGCTGTAAG 2819
Qy 721 TCCAAATACAGTCTCTTAATGCTGAACACCTGCACATCCGGAGTTCACCGGTCAAG 780
Db 2820 TCCAAATACAGTCTCTTAATGCTGAACACCTGCACATCCGGAGTTCACCGGTCAAG 2879
Qy 781 TGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGAGTTCAGCTTGGTCA 840
Db 2880 TGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGAGTTCAGCTTGGTCA 2939
Qy 841 CCAAAGACGGGACGCTTGTGCTTACGACATGAGAGTGCAGATCGGGGATGACCTGC 900
Db 2940 CCAAAGACGGGACGCTTGTGCTTACGACATGAGAGTGCAGATCGGGGATGACCTGC 2999
Qy 901 AGTGCACTGGCCCTGTATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGC 960
Db 3000 AGTGCACTGGCCCTGTATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGC 3059
Qy 961 TGGAGACAGGCCCTAACCTCGCCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT 1020
Db 3060 TGGAGACAGGCCCTAACCTCGCCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT 3119
Qy 1021 TCCTGCTCGGTGACAGTGTGCTTGAACACAGAGCTCAGCGTTCCAGGGGATYTG 1080
Db 3120 TCCTGCTCGGTGACAGTGTGCTTGAACACAGAGCTCAGCGTTCCAGGGGATYTG 3177
Qy 1081 NCAGCCCCCGGCTCAGCAGTGGAAACAGGCGCTCGNAGCAGNAGTNGGGG 1140
Db 3178 GCCAGCCCCCGGCTCA--CAGTGGGAACCAAGGCGCTCG-----CAGCAGCAAGGTGGG 3230
Qy 1141 CAAGCNAGAAATGCTCCAGGATTTACANCTGAGCCCTGAGCCCTGCTGAADA 1200
Db 3231 GCAAGCAAGATGCTCTCCAGGATTTACACTGAGCCCTGCCCCAC-----CTGCTGA 3284
Qy 1201 AAACAATTCCGCCACGTGAAGAGACAGAAGAGGATGAGGATTTNACCTYGGGGAA 1260
Db 3285 AAAAACAATCCGCCACGTGAAGAGACAGAAGAGGATGAGGATTTNACCTYGGGGAA 3340
Qy 1261 ACAAAAAGGATCTTTTCTGCCCCCTGCTCAGTNGAGTGGCCCTGNAACCGCTGG 1320
Db 3341 ACAAAAAGGATCTTTTCTGCCCCCTGCTCAGTNGAGTGGCCCTGNAACCGCTGG 3396
Qy 1321 ANTCAAGTACCATTTGTCAGANCAAGGAGAGCAGCTTCCAGCCTGGGTCAAGAGG 1380
Db 3397 GATCAGTGACCATTTGTTGGCAGA-CAGGGAGAGCAGCTTCCAGCCTGGGTCAAGAGG 3455
Qy 1381 GTGGGCGAGCCCTTGGCCCCCTCAACCTTNCAGGCTGCTGTGNAGAGTGTCAAGTGTGA 1440
Db 3456 GTGGGCGAGCCCTTGGCCCCCTCAACCTTNCAGGCTGCTGTGNAGAGTGTCAAGTGTGA 3513
Qy 1441 AGGNCACCAANTCAGNTTTCAGTGCAGAACCAAGTNCAGAGGATATGCCCCCAGNTA 1500
Db 3514 AGGCCCCAAATC---AGGTTTCAGTGCAGAACCAAGTNCAGAGGATATGCCCCCAGNTA 3569

Qy 1501 GGTAAANNNGGGGCCCTCTNAAACCCCTTCCCTNGGCTTNCACCTNGGCGAGCTCANCCC 1560
Db 3570 GTTAA-----GGGGGCCCTCTAAACCCCTTGCCTGGCCTCA---CCTGGCGAGCTCA-CCC 3621
Qy 1561 CTTTGGGTAGGGGAAAGAAATGCTGACCTTGGGAAGGCTWCCCTGCTAGAAATACAC 1620
Db 3622 CTTTGGGTAGGGGAAAGAAATGCTGACCTTGGGAAGGCT-CCCTGGTAGAAATACAC 3680
Qy 1621 CACACTTTTTCAGGTTGTTGCAACACAGGCTCCTGAGTTGACCTCTGGTTTCAGCCAGGACC 1680
Db 3681 CACACTTTTTCAGGTTGTTGCAACACAGGCTCCTGAGTTGACCTCTGGTTTCAGCCAGGACC 3740
Qy 1681 AAAGAAGGTGTGAAGTGAAGTGTCTCATGTCAGTNCCTCCAGACATGTGCCCCCTTTGCTGTG 1740
Db 3741 AAAGAAGGTGTGAAGTGAAGTGTCTCATGTTCTCAGT-CCCCAGACATGTGCCCCCTTTGCTGTG 3799
Qy 1741 GCTACCACTCTTCCCAAGAGCAGAGCCCGAGCCCTTTCAGGCCAGACATGCCCCAG 1800
Db 3800 GCTACCACTCTTCCCAAGAGCAGAGCCCGAGCCCTTTCAGGCCAGACATGCCCCAG 3859
Qy 1801 ACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAGAGTGAAGGTGATGAGGATATGCTGTA 1860
Db 3860 ACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAGAGTGAAGGTGATGAGGATATGCTGTA 3919
Qy 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGAGCY 1920
Db 3920 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGAGCY- 3976
Qy 1921 TCCGGCCCCCAGNCCCACTNATCAGTGTNCAGCGTGTCTGTTTCCCGAGCAGCAGC- 1980
Db 3977 TCCGGCCCCCAGCAGCCCACTATCAGTGT---CCAGCGTGTGTTTCCCGAGCAGCAGC- 4032
Qy 1981 GCTNCAAGCATCANCACTGACACTNCAACCTTNGCCCTGCCCCCTNGGCCAGAGGATCTG 2040
Db 4033 -----TCAGCTACACATGACACTCACCTGCTGCCCTGCCCTGG--CCAGAGGATCTG 4083
Qy 2041 CCGNAGCGCACTTTTGCACTNTGATGNACCTCAAAGCAGCTTTTCATGCTGTCCTCTTNG 2100
Db 4084 CCG-ACGGCACTTTTGCACT--CTGATGACCTCAAAGCAGCTTTTCATGCTGCTCTGCA 4140
Qy 2101 GCAGGNCAGGNCAGGNCAGTGCACANCTGTAGNAGCATANGCAAGCCAGAGATGG 2160
Db 4141 GGGCAGGCGCAGGCGAGTGCACACT-----GTAGGAGCATAGCAAGCCAGAGATGG 4190
Qy 2161 GGTGNAAGGANCACAGCTTTGAGCTGCCANCACTGATGATNCTCAAACCTCTTN 2220
Db 4191 GGTGAA--GGGACACAGCTTTGAGCTGTCCA-CATGATGTGACT-CCTCAAACCTCTT- 4245
Qy 2221 NCCAGNATTTCTCTAAGATAGCANCCCTTNNCCCCCATTTGCCCCAGCTTAGCCTCTTCT 2280
Db 4246 --CCAGATTTCTTAAGATAGCACCCCC--TTCCCCCATTTGCCCCAGCTTAGCCTCTTCT 4301
Qy 2281 CCAGGGAGCTANTCTAGGATCACTAGATGATTAATAGCTGTGNAATGTCAGGGGG 2340
Db 4302 CCAGGGAGCTA-CTCAGGACTCAGCTAGCATTAATAGCTGTG-AATCGTCAGGGGG 4359
Qy 2341 TGCTGCTAGCTCAACCTCTGGGCGAGGGAGCGCGAGACTCCGTTGGGAGAGCTCAT 2400
Db 4360 TGCTGCTAGCTCAACCTCTGGGCGAGGGAGCGCGAGACTCCGTTGGGAGAGCTCAT 4419
Qy 2401 TCCACATCTTTGCAAGACAGCCTTTTNGTCTCAGCTGTCTCCACATTTAGTTCAGACTGCTCCC 2460
Db 4420 TCCACATCTTTGCAAGACAGCCTTTT--GTCCAGCTGTCCACATTTAGTTCAGACTGCTCCC 4478
Qy 2461 GGGGAGAGCCCCCGCCCCCAGCAGACATAAAGAACTGCAGCTTTGGTACTGACAGTCTG 2520
Db 4479 GGGGAGAGCCCCCGCCCCCAGCAGACATAAAGAACTGCAGCTTTGGTACTGACAGTCTG 4538
Qy 2521 GGTGTGAGAACTCTTTGTAAGCAATAAAGTTTGGGTGATGACAAATGTTAAAAA 2578
Db 4539 GGTGTGAGAACTCTTTGTAAGCAATAAAGTTTGGGTGATGACAAATGTTAAAAA 4596


```
RESULT 4
US-10-283-975A-294
; Sequence 294, Application US/10283975A
; Publication No. US20040110792A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
; FILE REFERENCE: CDS 293 PCT
; CURRENT APPLICATION NUMBER: US/10/283,975A
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/340,938
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/338,997
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/340,081
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/341,012
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 294
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: HUMAN
US-10-283-975A-294

Query Match      70.3%; Score 1850.2; DB 19; Length 4596;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

QY      1  CCCCTCTCACAGCCAGCCAGGCGCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
DB      2100  CCCCTCTCACAGCCAGCCAGGCGCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 2159

QY      61  CTGACGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120
DB      2160  CTGACGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 2219

QY      121  GCCCTTGGAGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC 180
DB      2220  GCCCTTGGAGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC 2279

QY      181  ACCAGACCTTCATGCCAGCCAGCGAGAGCTTTCCGCAAGGGCCCGAGGCCCGGCCAG 240
DB      2280  ACCAGACCTTCATGCCAGCCAGCGAGAGCTTTCCGCAAGGGCCCGAGGCCCGGCCAG 2339

QY      241  CTGAGGAGACAACAGGCGAGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCCCAG 300
DB      2340  CTGAGGAGACAACAGGCGAGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCCCAG 2399

QY      301  AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGGTCTGGGATGTGGNACCT 360
DB      2400  AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGGTCTGGGATGTGGNACCT 2459

QY      361  TACCTCTGTCTCTCCCTGGAGCGAGCCCTTCGCAAGAACCCAGCTCACAGAGCGGAAAG 420
DB      2460  TACCTCTGTCTCTCTGGAGCGAGCCCTTCGCAAGAACCCAGCTCACAGAGCGGAAAG 2519

QY      421  CAACCGTCCCGAGCAGGAACCTGACAGAGCTGGAAATAGAAATTAATCTCAACAGCCTGT 480
DB      2520  CAACCGTCCCGAGCAGGAACCTGACAGAGCTGGAAATAGAAATTAATCTCAACAGCCTGT 2579

QY      481  CCCAGCCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTCGTCTGCTCTAGCATCCAGCC 540
DB      2580  CCCAGCCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTCGTCTGCTCTAGCATCCAGCC 2639

QY      541  TCTCCCTGTGCGATGACAGTGAAGAAACCCATCAAAAGCCCTCTCAAAAGCTCGCGGGACA 600
DB      2640  TCTCCCTGTGCGATGACAGTGAAGAAACCCATCAAAAGCCCTCTCAAAAGCTCGCGGGACA 2699

QY      601  CCCTGAGCTCAGGCGTACACTCTGGAGCAGCAGCGGCGAGGCTCGAAAGCTCCAGCTGGA 660
DB      3741  AAAGAAGGTGTGAAGTGAAGTGGTGTCTCAGT - CCCCAGACATGTGCCCTTTTGTGCTG 3799
```

```
DB      2700  CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCCGAGGCTCGAAGCTCCAGCTGA 2759
QY      661  ACATGGTCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGGTGTGAAG 720
DB      2760  ACATGGTCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGGTGTGAAG 2819
QY      721  TCCAAATACAGTCTCTTAATGGTGAACACCTTGCATCCGGAGGTTCACCGGGTCAAG 780
DB      2820  TCCAAATACAGTCTCTTAATGGTGAACACCTTGCATCCGGAGGTTCACCGGGTCAAG 2879
QY      781  TGGGAGACATCGCCACTGGCATTCAGCAGCCAGATCCAGCTGCGAGCTTCAGCTTGGTCA 840
DB      2880  TGGGAGACATCGCCACTGGCATTCAGCAGCCAGATCCAGCTGCGAGCTTCAGCTTGGTCA 2939
QY      841  CAAAGACGGGACGCTTTCGCTACGACATGAGAGGTGCCAGACTCGGGGCATTCGACCTGC 900
DB      2940  CAAAGACGGGACGCTTTCGCTACGACATGAGAGGTGCCAGACTCGGGGCATTCGACCTGC 2999
QY      901  AGTGACACTGCGCCCTGATGCGAGCTTCGCTCGAGCTGGAGGTCAAGCATGSCCAGC 960
DB      3000  AGTGACACTGCGCCCTGATGCGAGCTTCGCTCGAGCTGGAGGTCAAGCATGSCCAGC 3059
QY      961  TGGAGAACAGGCGCCTTAACCTCGCCCTCCACCGCGGCTCCACACTGCGCGGAAAGCAGCCT 1020
DB      3060  TGGAGAACAGGCGCCTTAACCTCGCCCTCCACCGCGGCTCCACACTGCGCGGAAAGCAGCCT 3119
QY      1021  TCCTGCTCGGTGACGATGCTGCCCTGAAACACAGAGCTCAGCGTTCAGCGTTCAGAGGGAT 1080
DB      3120  TCCTGCTCGGTGACGATGCTGCCCTGAAACACAGAGCTCAGCGTTCAGCGTTCAGAGGGAT 3177
QY      1081  NCCAGCCCCCGGCTCARGCAGTGGGAACCAAGGCGCTCGCNCAGCAGNAAAGGTGGGG 1140
DB      3178  GCGAGCCCCCGGCTCA--CAGTGGGAACCAAGGCGCTCG-----CAGCAGCAAGGTGGGG 3230
QY      1141  CAAGCAGAAATGCTCCCGAGGATTTACANCTTGAGCCCNMTGCCCNACCTGCTGTAADA 1200
DB      3231  GCAAGCAGAAATGCTCCCGAGGATTTACACTGAGCCCTGCCCCAC-----CCTGCTGA 3284
QY      1201  AAACAYTNCGCCACGCTGAAGAGACAGAGAGGATGNCAGAGGTNNACCTTGGGGAA 1260
DB      3285  AAACAACTCCGCCACGCTGAAGAGACAGAGAGGATGNCAGAGGTNNACCTTGGGGAA 3340
QY      1261  AAAAAACAGGATCTTTTCTGCTCCCTGCTCCAGTNCAGGTTCGCTGNCACCTGCTGG 1320
DB      3341  AAAAAACAGGATCTTTTCTGCTCCCTGCTCCAGTNCAGGTTCGCTGNCACCTGCTGG 3396
QY      1321  ANTCACTGACATTTGTCAGANACAGGGAGAGAGCTTCAGCTTGGGTCAAGAGGG 1380
DB      3397  GATCAGTGACATTTGTTGGCAGA-CAGGGGAGAGAGCTTCAGCTTGGGTCAAGAGGG 3455
QY      1381  GTGGGCGAGCCCTTCGGCCCTCACCTNCAGGCTGCTGTGNAGAGTGTCAAGTGTGA 1440
DB      3456  GTGGGCGAGCCCTTCGGCCCTCACCTNCAGGCTGCTGTGNAGAGTGTCAAGTGTGA 3513
QY      1441  AGGNCNCAAACTCAGNTTTCAGTGCAGAACACAGTNCAGCAGGTATGCCCGCCCGNTA 1500
DB      3514  AGGNCNCAAACTC---AGGTTCACTGCAGAACACAGT- CAGCAGGTATGCCCGCCCGTAG 3569
QY      1501  GGTTAANNVGGGGCCCTCTTAAACCCCTTGCTNGGCTTCACCTNGGCCAGCTCANCCC 1560
DB      3570  GTTAA-----GGGGGCCCTCTAAACCCCTTGCTGGCCCTCA---CCTGGCCAGCTCA-CCC 3621
QY      1561  CTTTTGGGTGTAGGGGAAAGAAATGCTGACCCCTGGGAAAGCTTCCCTGGTGTAGATACAC 1620
DB      3622  CTTTTGGGTGTAGGGGAAAGAAATGCTGACCCCTGGGAAAGCT-CCCTGGTGTAGATACAC 3680
QY      1621  CACACTTTTTCAGGTGTGTGCAACACAGGTCTGAGTTGACCTCTGTGTTTCCAGCCAGGACC 1680
DB      3681  CACACTTTTTCAGGTGTGTGCAACACAGGTCTGAGTTGACCTCTGTGTTTCCAGCCAGGACC 3740
QY      1681  AAAGAAGGTGTGAAGTGAAGTGGTTCCTCAGTNCCTCCAGACATGTGCCCTTTGCTGCTG 1740
DB      3741  AAAGAAGGTGTGAAGTGAAGTGGTTCCTCAGT - CCCCAGACATGTGCCCTTTTGTGCTG 3799
```

QY 1741 GCTACCACTCTTCCCGAGAGCAGAGGCCCCGAGCCCTTTCAGGCCAGCACTGCCCCAG 1800
DB GCTACCACTCTTCCCGAGAGCAGAGGCCCCGAGCCCTTTCAGGCCAGCACTGCCCCAG 3859
QY 1801 ACTGCTGGCACTCAGTTCCCTCATCTCTAAGGTGAAGGTGATGCAAGGATGCTCTGA 1860
DB ACTGCTGGCACTCAGTTCCCTCATCTCTAAGGTGAAGGTGATGCAAGGATGCTCTGA 3919
QY 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAGGNAAGCAGCAGAGAGAGAGY 1920
DB CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAGGNAAGCAGCAGAGAGAGAGY 3976
QY 1921 TCCGGCGCCAGNCCCACTNATCAGTGTGNCAGCGTGTCTGCTGCTGCTGCTGCTGCTG 1980
DB TCCGGCGCCAGNCCCACTNATCAGTGTGNCAGCGTGTCTGCTGCTGCTGCTGCTGCTG 4032
QY 1981 GCTNAGNCACTCANCACACTGACACTNCACCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
DB TCCGGCGCCAGNCCCACTNATCAGTGTGNCAGCGTGTCTGCTGCTGCTGCTGCTGCTG 4083
QY 2041 CCGNACGCACTTTGCACTGATGACACTTCAAGCACTTTTCACTGCTGCTGCTGCTGCTG 2100
DB CCG-ACGCACTTTGCACT--CTGATGACCTCAAGCACTTTTCACTGCTGCTGCTGCTG 4140
QY 2101 GCAGGNCAGGNCAGGNCAGTGCANCTGTAGNAGCATPANGCAANGCCAGAGATGG 2160
DB GGGCAGGCGAGGCGAGTGACACT-----GTAGGAGCATAGCAAGCCAGGAGATGG 4190
QY 2161 GGTGNAAGGANCACAGCTTTGAGCTGTCANCACTGATGATGACTNCTCAAACTCTTN 2220
DB GGTGAA--GGGACACAGCTTTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4245
QY 2221 NCCAGNATTTCTTAAGATAAGCAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
DB CCAGATTTCTTAAGNATAGCAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4301
QY 2281 CCCAGGAGGATNCTCAGGACTCAGCTAGCATTAATCAGCTGTGNAATCGTCAGGCGG 2340
DB CCCAGGAGGACTA-CTCAGGACTCAGCTAGCATTAATCAGCTGTG-AATCGTCAGGCGG 4359
QY 2341 TGTCTGTAGCTCAACTCTCGGGCAGGAGCGCGAGACTCGGTGGGAGAGCTCAT 2400
DB TGTCTGTAGCTCAACTCTCGGGCAGGAGCGCGAGACTCGGTGGGAGAGCTCAT 4419
QY 2401 TCCACATCTTGCAGACAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
DB TCCACATCTTGCAGACAGCTTTT-GTCCAGCTGTCCCATTTGAGTTCAGCTGCTGCTG 4478
QY 2461 GGGGAGAGAGCCCCGCGCCCGCAGCACATAAAGAACTGCGAGCTTTGTAAGTGTG 2520
DB GGGGAGAGAGCCCCGCGCCCGCAGCACATAAAGAACTGCGAGCTTTGTAAGTGTG 4538
QY 2521 GGTGTAGAGAACTCTTTGTAAGCAATAAGTTTGGGTGTATGACAAATGTTAAAAA 2578
DB GGTGTAGAGAACTCTTTGTAAGCAATAAGTTTGGGTGTATGACAAATGTTAAAAA 4596

RESULT 5

US-10-087-192-887
; Sequence 887, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: CANCER
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087.192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586

; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 887
; LENGTH: 4684
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-087-192-887

Query Match 67.8%; Score 1783.4; DB 13; Length 4684;

Best Local Similarity 89.3%; Pred. No. 0;

Matches 2297; Conservative 5; Mismatches 193; Indels 87; Gaps 35;

QY 1 CCCCTCTCAGAGCCAGGCCATCCAGAGGGCTTGAGAAAGAGCCCATCCACCGCTGT 60
DB CCCCTCTCAGAGCCAGGCCATCCAGAGGGCTTGAGAAAGAGCCCATCCACCGCTGT 2209
QY 61 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120
DB CTGAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 2269
QY 121 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCCAATTACC 180
DB GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCCAATTACC 2329
QY 181 ACCAGACCTTCCATGCCAGCCGAGAGAGCTTTCCCAAGGGCCCCAGGGCCCCGCGCAG 240
DB ACCAGACCTTCCATGCCAGCCGAGAGAGCTTTCCCAAGGGCCCCAGGGCCCCGCGCAG 2389
QY 241 CTGAGGAGACACAGGCGAGAGCCCTTAAGTCTCAGCCCTCTCTCCACAGAGCCCCCAG 300
DB CTGAGGAGACACAGGCGAGAGCCCTTAAGTCTCAGCCCTCTCTCTCCACAGAGCCCCCAG 2449
QY 301 AGCCAAACAAAGTCTCTCCCTTGAAGTCTGAGCAAGGAGGAGTCTGGGATGTGGAAACCT 360
DB AGCCAAACAAAGTCTCTCCCTTGAAGTCTGAGCAAGGAGGAGTCTGGGATGTGGAAACCT 2509
QY 361 TACCTCTGTCTCTCTCGAGCCAGCCCTTGCAGAAACCCAGCTCACCAGAGCGGAAAG 420
DB TACCTCTGTCTCTCTCGAGCCAGCCCTTGCAGAAACCCAGCTCACCAGAGCGGAAAG 2569
QY 421 CAACCGTCCGAGAGCAGGAACTGAGCAGCTGGAATAGAAATATTCTCTCAACAGCTGT 480
DB CAACCGTCCGAGAGCAGGAACTGAGCAGCTGGAATAGAAATATTCTCTCAACAGCTGT 2629
QY 481 CCAGCCATTTCTCTGAGGAGCAGAGCAAAATTTCTCTGCTGCTCAGCATCAGACGCC 540
DB CCAGCCATTTCTCTGAGGAGCAGAGCAAAATTTCTCTGCTGCTCAGCATCAGACGCC 2689
QY 541 TCTCCCTGTCTCGGATGACAGTGAGAAAGAACCCATCAAAAGGCTCTCAAAAGCTCGCGGACA 600
DB TCTCCCTGTCTCGGATGACAGTGAGAAAGAACCCATCAAAAGGCTCTCAAAAGCTCGCGGACA 2749
QY 601 CCCTGAGCTCAGCGGTACACTCTGAGAGAGCAGAGCCGAGGCTCGAAGCTCCAGCTGGA 660
DB CCCTGAGCTCAGCGGTACACTCTGAGAGAGCAGAGCCGAGGCTCGAAGCTCCAGCTGGA 2809
QY 661 ACATGCTGTGGCCGGGGGGCCACCGACACCCCAAGCTATTTCATGCTGTGAAG 720
DB ACATGCTGTGGCCGGGGGGCCACCGACACCCCAAGCTATTTCATGCTGTGAAG 2869
QY 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCACCGGCTCAAG 780
DB TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCACCGGCTCAAG 2929
QY 781 TGGGAGACATCGGCATCTGGCATCAGAGCCAGATCCAGCTGACGCTTACGCTTGTGCA 840
DB TGGGAGACATCGGCATCTGGCATCAGAGCCAGATCCAGCTGACGCTTACGCTTGTGCA 2989
QY 841 CCAAAGAGCGGCGACCTGTTGCTACGATGAGGTTCCAGACTCGGGCATCCACCTGC 900
DB CCAAAGAGCGGCGACCTGTTGCTACGATGAGGTTCCAGACTCGGGCATCCACCTGC 3049


```
QY 901 AGTGACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGAGGGTCAAGCATGGCCAGC 960
Db 3050 AGTGACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGAGGGTCAAGCATGGCCAGC 3109
QY 961 TGGAAACAGGCTTAACCTGCTCCCTCAACCGCCGGCTCCACACTGCGGAAAGCAGCCT 1020
Db 3110 TGGAAACAGGCTTAACCTGCTCCCTCAACCGCCGGCTCCACACTGCGGAAAGCAGCCT 3168
QY 1021 TCCTGCTGGTGACAGTGTCTGCTGAAACACAGGCTCAGCGTTCCTCCAGGGATYTG 1080
Db 3169 TCCTGCTGGTGACAGTGTCTGCTGAAACACAGGCTCAGCGTTCCTCCAGGG---ATC 3224
QY 1081 NCCAGCCCCCGGCTCARGAGTGGGAACACAGGCGCTCGNAGCAGGAGTTCCTCCAGGG 1140
Db 3225 TGCCAGCCCCCGGCTCARGAGTGGGAACACAGGCGCTCGNAGCAGGAGTTCCTCCAGGG 3277
QY 1141 CAAGNAGATGCTCCCAAGGATTTTCANCTGAGCCGCTGCCCCCAGCCCTGCTGAADA 1200
Db 3278 GCAAGCAGATGCTCCCAAGGATTTTCANCTGAGCCGCTGCCCCCAGCCCTGCTGAADA 3334
QY 1201 AAACAYTNCGCCACGCTGAGAGACAGAGGAGGATGNCAGGAGTTCNACCTYGGGAA 1260
Db 3335 AAACACT---CCGCCACGCTGAAGAGACA---GAGGAGGATGCGAGGATTCACCTCGGAAA 3388
QY 1261 ACAAACAGGGATCTTNTCTGCCCCCTGCTCCAGTNGAGTGGCTGNCACCCGCTTGG 1320
Db 3389 CAACAGGATCTTCTGCTGCCCCCTGCTCCAG-----TCAGTTGGCTGACCCGCTTGG- 3440
QY 1321 ANTCAGTGACCATTTGTTGGCAGANAGGGGAGAGAGCTTCCAGCCTGGGTGAGAGGG 1380
Db 3441 -ATCAGTGACCATTTGTTGGCAGAGAGAGAGCTTCCAGCCTGGGTGAGAGGG 3498
QY 1381 GTGGGCGAGCCCTCGGCCCTCACCCTNCAGGCTGCTGTCNAGAGTGTCAAGTGTA 1440
Db 3499 GTGGGCGAGCCCTCGGCCCTCACCCT---CCAGGCTGCTGTG---AGAGTGTCAAGTGTA 3556
QY 1441 AGGNCNCCAAANCTCAGGNTTCAGTGCAGAACAGGTCNAGCAGGATGTCGCCGCCGNTA 1500
Db 3557 AGGCCCAACTC---AGGTTCAAGTGAGAACAGGTCNAGCAGGATGTCGCCGCCGNTA 3612
QY 1501 GGTAAAGGGGGCCCTCTNAAACCCCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
Db 3613 GTTAA---GGGGGCCCTCTAAACCCCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3664
QY 1561 CTTTGGGTGAGGGGAAAGATGCTGACCTCGGAGGCTTCCCTGCTGAGATACAC 1620
Db 3665 CTTTGGGTGAGGGGAAAGATGCTGACCTCGGAGGCTTCCCTGCTGAGATACAC 3723
QY 1621 CACACTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCTGCTGAGCCAGGACC 1680
Db 3724 CACACTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCTGCTGAGCCAGGACC 3783
QY 1681 AAAGAAGGTGTGAAGTGTGTTCTCAGTNCNCCAGACATGTCGCCCTTGTGCTG 1740
Db 3784 AAAGAAGGTGTGAAGTGTGTTCTCAGTNCNCCAGACATGTCGCCCTTGTGCTGCTG 3842
QY 1741 GCTACCACTTTCCCGAGAGCAGCGCCCGGAGCCCTTTCAGGCCCCAGCACTGCCCCAG 1800
Db 3843 GCTACCACTTTCCCGAGAGCAGCGCCCGGAGCCCTTTCAGGCCCCAGCACTGCCCCAG 3902
QY 1801 ACTCGCTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGTTGATGAGGATATGCTGTA 1860
Db 3903 ACTCGCTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGTTGATGAGGATATGCTGTA 3962
QY 1861 CAGGAACAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920
Db 3963 CAGGAACAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4020
QY 1921 TCCGGGCCCCAGNCCCACTNATCAGTGTGTCNAGGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Db 4021 TCCGGGCCCCAGNCCCACTNATCAGTGTGTCNAGGCTGCTGCTGCTGCTGCTGCTGCTG 4069
QY 1981 GCTNCAGNATCANCACTGACATNCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
```

```
Db 4070 GCACAGCTCAGCATCACACTGACACTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4127
QY 2041 CCNNAACGCACTTTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
Db 4128 CCG-ACGCACTTTGCACT-TCGATG-ACCTCAAGACATTTTCATGCTGCTGCTGCTGCTG 4184
QY 2101 GCAGGNCAGGNCAGGNCAGTGCACACTGTTAGNAGCATANGCAAGCAGGAGATGG 2160
Db 4185 GGGCAGGCGAGGCGAGTGACACTGT-----AGGAGCATAGCAAGCAGGAGATGG 4234
QY 2161 GGTGNAAGGAGNACAGCTTTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
Db 4235 GGTGAA--GGGACACAGCTTTGAGCTGTCCA-CATGATGTGACT-CCTCAACCTCTT- 4289
QY 2221 NCCAGNATTTCTCTAAGAAATAGCAGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Db 4290 --CCAGATTTCTCTAAGAAATAGCAGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCT 4345
QY 2281 CCCAGGGAGGCTAGCTCAGGACTCACTGAGCACTAAATCAGCTGTGNAATCGTCAGGGGG 2340
Db 4346 CCCAGGGAGGCTA-CTCAGGACTCAGCTAGCACTAAATCAGCTGTG-ATCGTCAGGGGG 4403
QY 2341 TGTCTGCTAGCTCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Db 4404 TGTCTGCTAGCTCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4463
QY 2401 TCCACATCTTCCAAAGACAGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Db 4464 TCCACATCTTCCAAAGACAGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4522
QY 2461 GGGGAGAGAGCCCGGCCCCCAGCAGCATAAAGAACTGAGCTGCTGCTGCTGCTGCTGCTG 2520
Db 4523 GGGGAGAGAGCCCGGCCCCCAGCAGCATAAAGAACTGAGCTGCTGCTGCTGCTGCTGCTG 4582
QY 2521 GGTGTGAGAACTCTTTGTAGCAATAAAGTTTGGGCTGATGACAAATGTT 2572
Db 4583 GGTGTGAGAACTCTTTGTAGCAATAAAGTTTGGGCTGATGACAAATGTT 4634
```

RESULT 6

US-09-871-889-2

; Sequence 2, Application US/09871889

; Patent No. US20020042499A1

; GENERAL INFORMATION:

; APPLICANT: Greene, Warner C.

; APPLICANT: Lin, Xin

; APPLICANT: Gelezuinas, Romas

; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED

; FILE REFERENCE: 30448.61USU1

; CURRENT APPLICATION NUMBER: US/09/871,889

; PRIOR FILING DATE: 2001-06-01

; PRIOR APPLICATION NUMBER: 09/257,703

; PRIOR FILING DATE: 1999-02-25

; PRIOR APPLICATION NUMBER: 60/076,299

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 2844

; TYPE: DNA

; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)

US-09-871-889-2

Query Match 37.1%; Score 977; DB 9; Length 2844;

Best Local Similarity 100.0%; Pred. No. 1.7e-290; Indels 0; Gaps 0;

Matches 977; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCTCTCAAGCCAGGCGCATCCAGAGGGGCTGAGGAAAGAGCCATCCACCGCTGT 60

Db 1868 CCCCTCTCAAGCCAGGCGCATCCAGAGGGGCTGAGGAAAGAGCCATCCACCGCTGT 1927

```
QY 61 CTGACGCGAGCTGGAGGGAAGGTGAACCGGCACCTACAGCAAGTGGAGGTCTGAAGA 120
Db 1928 CTGACGCGAGCTGGAGGGAAGGTGAACCGGCACCTACAGCAAGTGGAGGTCTGAAGA 1987
QY 121 GCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 180
Db 1988 GCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 2047
QY 181 ACCAGACCTTCATGCCCCAGCGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCGGCCAG 240
Db 2048 ACCAGACCTTCATGCCCCAGCGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCGGCCAG 2107
QY 241 CTGAGGAGACAAAGGAGAGCCCTTAAGCTTCAGCTCCAGCTTCCTCCACAGAGCCCCCAG 300
Db 2108 CTGAGGAGACAAAGGAGAGCCCTTAAGCTTCAGCTTCAGCTTCCTCCACAGAGCCCCCAG 2167
QY 301 AGCCAAACAAGTCTCTCCCTTGACTTTTGACCAAGGAGGAGTCTGGGATGTGGCAACCT 360
Db 2168 AGCCAAACAAGTCTCTCCCTTGACTTTTGACCAAGGAGGAGTCTGGGATGTGGCAACCT 2227
QY 361 TACCTCTGTCTCTCTGGAGCCAGCCCTTGCCAGAAACCCAGCTCCACAGAGCGGAAAG 420
Db 2228 TACCTCTGTCTCTCTGGAGCCAGCCCTTGCCAGAAACCCAGCTCCACAGAGCGGAAAG 2287
QY 421 CAACCGTCCCGAGCAGGAATCTGAGCAGAGCTGGAAATAGAAATATTCTCAACAGCCTGT 480
Db 2288 CAACCGTCCCGAGCAGGAATCTGAGCAGAGCTGGAAATAGAAATATTCTCAACAGCCTGT 2347
QY 481 CCCAGCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATGCACAGCC 540
Db 2348 CCCAGCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATGCACAGCC 2407
QY 541 TCTCCCTGTGCGATGACAGTGAAGAAACCCATCAAAAGGCTCTCAAAAGCTTCGCGGACA 600
Db 2408 TCTCCCTGTGCGATGACAGTGAAGAAACCCATCAAAAGGCTCTCAAAAGCTTCGCGGACA 2467
QY 601 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGCGGAGCTCGAAGCTCCAGCTGGA 660
Db 2468 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGCGGAGCTCGAAGCTCCAGCTGGA 2527
QY 661 ACATGCTGCTGGCCGGGGGGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 720
Db 2528 ACATGCTGCTGGCCGGGGGGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 2587
QY 721 TCCAAATACAGTCTCTTAATGGTGAAACCTTGACATCCCGGAGTTCACCGGCTCAAAG 780
Db 2588 TCCAAATACAGTCTCTTAATGGTGAAACCTTGACATCCCGGAGTTCACCGGCTCAAAG 2647
QY 781 TGGAGACATCGCACCTGGCATCAGCAGCCAGATCCCGAGCTGCAGCTTCAGCTTGGTCA 840
Db 2648 TGGAGACATCGCACCTGGCATCAGCAGCCAGATCCCGAGCTGCAGCTTCAGCTTGGTCA 2707
QY 841 CCAAGACGGCAGCCCTTTCCTTACACATGGAGGTCCAGACTCGGGATCGACCTGC 900
Db 2708 CCAAGACGGCAGCCCTTTCCTTACACATGGAGGTTCAGACTCGGGATCGACCTGC 2767
QY 901 AGTGCACACTGGCCCTGATGGCAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 960
Db 2768 AGTGCACACTGGCCCTGATGGCAGCTTCGCTCGAGCTGGAGGTTCAGCATGGCCAGC 2827
QY 961 TGGAGAACAGCCCTAA 977
Db 2828 TGGAGAACAGCCCTAA 2844
```

RESULT 7

```
US-10-087-192-886
; Sequence 886, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
```

```
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 886
; LENGTH: 73967
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-886
```

```
Query Match 37.08; Score 973.4; DB 13; Length 73967;
Best Local Similarity 84.4%; Pred. No. 7.6e-289;
Matches 1487; Conservative 5; Mismatches 183; Indels 87; Gaps 35;

QY 811 AGATCCCGAGCTGCAGCCTTCAGCTTGGTCCACCAAGACGGCAGCCTGTTCGCTACGACA 870
Db 62243 AGATCCCGAGCTGCAGCCTTCAGCTTGGTCCACCAAGACGGCAGCCTGTTCGCTACGACA 62302
QY 871 TGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACTGGCCCCCTGATGGCAGCTTCG 930
Db 62303 TGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACTGGCCCCCTGATGGCAGCTTCG 62362
QY 931 CCTGGAGCTGGAGGTTCAAGCATGGCCAGCTGGAGAACAGGCCCTAACCCCTGCCCTCCAC 990
Db 62363 CCTGGAGCTGGAGGTTCAAGCATGGCCAGCTGGAGAACAGGCCCTAACCCCTGCCCTCCAC 62422
QY 991 CGCGGGCTCCACACTGCGGAAAGACAGCCTTCTCTGCTCGGTGCACGATGCTGCCCTGAAA 1050
Db 62423 CGCGGGCTCCACACTGCGCGG--AAGCAGCTTCTCTGCTGGTGACAGATGCTGCCCTG-AA 62480
QY 1051 ACACAGCTCAGCGTTTCCAGGGGATYTNCCAGCCCCCGGCTCARGANTGGGAACC 1110
Db 62481 ACACAGCTCAGCGTTTCCAGGGG---ATCTGCCAGCCCCCGGCTCAGCAGTGGGACCAG 62537
QY 1111 AGGCGCTCGNAGCNAGNAGGTTGGGGGCAAGCNAGATGCTTCCAGGATTTACAN 1170
Db 62538 GGCTTCGACAGCAAG--GTGGGGGCAAGCAGAAATGCTTCCAGGATTTACAC 62590
QY 1171 CCTGAGCCTGTCGCCCCCAGCCTCTGTAADAAAACAYTNCCGCGCACGTAAGAGACAGAG 1230
Db 62591 C---TGAGCCCTGCCACCCTCTGTAAGAAAACACT--CGCGCACGTAAGAGAGACA---- 62641
QY 1231 GAGGATGNCAGGAGTTNNACCTYGGGGAACAAAACAGGGATCTTTTCTGCCCCCTGC 1290
Db 62642 GAGGAGGATGGCAGGAGTTTACCTCGGGAACAAACAGGATCTTCTCTGCCCTGCTCCAG- 62700
QY 1291 TCAGTNCAGATTTGGCTGNACCCGCTTGGANTCAGTGACATTTGTTGGCAGANCAAGGG 1350
Db 62701 -----TCGAGTTTGGCCTGA--CCCGCTTGGATCAGTGACCATTTTGTGTGGCAGA-CAGGG 62751
QY 1351 GAGAGCAGCTTCCAGCTGGGTGAGAGGGGTGGGAGCCCTTCGCCCCCTCACCCCTNC 1410
Db 62752 GAGAGCAGCTTCCAGCTGGGTGAGAGGGGTGGGAGCCCTTCGCCCCCTCACCCCT-C 62810
QY 1411 CAGGCTGCTGAGAGTGTCAAGTGTGTAAGGNCACAAANCTCAGGNTTTCAGTGCGAGA 1470
Db 62811 CAGGCTGCTGTG-AGAGTGTCAAGTGTGTAAGGGGCCAAATC---AGGTTTCAGTGCGAGA 62866
QY 1471 ACCAGTNCAGCAGGATGTCGCCCGCCGNTAGGTTAAANNNGGGGCCCTCTNAAACCCCTTG 1530
Db 62867 ACCAGGT-CAGCAGGATGTCGCCCGCCGNTAGGTTAA-----GGGGGCCCTCTTAAACCCCTTG 62921
QY 1531 CCTNGGCTNCACCTTNGGCGCAGCTCANCCCTTTTGGGTGTAGGGGAAAGATGCTGA 1590
Db 62922 CCTGGCCTCA---CCTGGCCAGCTCA-CCTCTTTTGGGTGTAGGGGAAAGATGCTGA 62977
QY 1591 CCCTGGGAAGGCTWCCTCTGTGTAAGATACACACACTTTTTCAGGTTTGTGCAACACAGGTC 1650
```

Db 62978 CCCTGGGAAGGCT-CCCTGGTAGAATACACACACTTTTCAGGTGTTGCAACACAGGTC 63036
QY 1651 CTGAGTTGACCTCTGGTTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGTCTCA 1710
Db 63037 CTGAGTTGACCTCTGGTTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGTCTCA 63096
QY 1711 GTNCCCCAGACATGTGCCCTTTGCTGTGCTGTACACTCTTCCCCAGAGCAGAGGCC 1770
Db 63097 GT-CCCCAGACATGTGCCCTTTGCTGTGCTGTACACTCTTCCCCAGAGCAGAGGCC 63155
QY 1771 CGAGCCCTTTCAGGCCACAGCACTGCCAGACTCGCTGCGACTCAGTTTCCCTCATCTGTA 1830
Db 63156 CGAGCCCTTTCAGGCCACAGCACTGCCAGACTCGCTGCGACTCAGTTTCCCTCATCTGTA 63215
QY 1831 AAGGTGAAGGTGATGCAAGATATGCCCTGACAGGAACAGTCTGTGGATGGACATGATCAG 1890
Db 63216 AAGGTGAAGGTGATGCAAGATATGCCCTGACAGGAACAGTCTGTGGATGGACATGATCAG 63275
QY 1891 TGCTNAAGNAAGCAGCAGAGAGAGCGYTCCGGCGCCCGAGNCCCACTNATCAGTGT 1950
Db 63276 TGCTAA--GGAAAGCAGCAGAGAGAGCGTCCGGCGCCCGAGNCCCACTNATCAGTGT-- 63331
QY 1951 NCCAGCGTCTGCTGTTNCCCCAGNAGCAGAGCTNCAAGNCACTGACATNCAACC 2010
Db 63332 -CCAGCGTCTGTTTCCC-----CAGAGCAGACTCAGCATCACACTGACACTCACC 63382
QY 2011 TNGCCCTGCCCTNCGCCANGAGGTACTCGGNAGCGCACTTTGCACTNCTGATGNACC 2070
Db 63383 CTGCGCTGCCCTG--CCAGAGGTGACTGCGC-ACGGCACTTTGAC-CTGTGATG-ACC 63437
QY 2071 TCAAGCACTTTTCATGGCTNCGCTCTNNGCAGGNCAGGNCAGGNCAGTGCANCT 2130
Db 63438 TCAAGCACTTTTCATGGCTGCGCTCTGCGAGGCGAGGCGAGTGCACACT----- 63490
QY 2131 GTAGNAGCATANGCAANGCCAGGAGATGGGGTGNAGGGANCAAGTCTTGAGTGTCC 2190
Db 63491 --GTAGGAGCATAGCAAGCCAGGAGATGGGGTGAA--GGGACACAGTCTTGAGTGTCC 63545
QY 2191 ANCATGATGTGACTNCTCAAACTCTNCCAGNATTTCTTAAGAAATAGCANCCTCC 2250
Db 63546 A-CATGCACTGTGACT-CTCAAACTCTT-CCAGATTTCTTAAGAAATAGCANCCTCC- 63599
QY 2251 TTNCCCAATGGCCAGCTTAGCTCTTCTCCAGGGAGCTTANCTCAGGACTCAGTAG 2310
Db 63600 -TTCCCAATGGCCAGCTTAGCTCTTCTCCAGGGAGCTA-CTCAGGACTCAGTAG 63657
QY 2311 CATTAATCAGCTGTGNAATCTCAGGGGTGTCTGAGCTCAACCTCTCTGGGGCAGG 2370
Db 63658 CATTAATCAGCTGTG-AATCGTCAGGGGTGTCTGCTAGCTCAACCTCTCTGGGGCAGG 63716
QY 2371 GGACCGGAGACTCCGTGGGAGAGCTCATTTCCACATCTTCCAGACAGCCTTTNGTC 2430
Db 63717 GGACCGGAGACTCCGTGGGAGAGCTCATTTCCACATCTTCCAGACAGCCTTT-GTG 63775
QY 2431 CAGCTGTCCAAATTGAGTCAGACTGTCTCCGGGGAGAGAGCCCGGCCCCAGCATAA 2490
Db 63776 CAGCTGTCCAAATTGAGTCAGACTGTCTCCGGGGAGAGAGCCCGGCCCCAGCATAA 63835
QY 2491 AGAACTGACGCTTGGTACTGAGAGTCTGGGTGTGAGAACTCTTTGTAAAGCAATAA 2550
Db 63836 AGAACTGACGCTTGGTACTGAGAGTCTGGGTGTGAGAACTCTTTGTAAAGCAATAA 63895
QY 2551 GTTTGGGGTGATGACAAATGTT 2572
Db 63896 GTTTGGGGTGATGACAAATGTT 63917

RESULT 8
US-10-296-115-97
; Sequence 97, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:

; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1e1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296.115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 97
; LENGTH: 1385
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-296-115-97

Query Match 28.7%; Score 754.2; DB 18; Length 1385;
Best Local Similarity 83.5%; Pred. No. 8.6e-22;
Matches 1205; Conservative 4; Mismatches 159; Indels 75; Gaps 31;

QY 1137 GGGCAACGACGATGCTCCAGGATTTACANCTTGAGCCNTGCCCANCCCTGCTG 1196
Db 15 GGGGCAAGCAAGATGCTCCAGGATTTACACC---TGAGCCCTGCCCACTGCTG 71
QY 1197 AADAAACAYTNCGCCCACTGTAAGAGACAGAAAGAGGATGNCAGGAGTTNNACCTYGG 1256
Db 72 AGAAAACT--CCGCCACGTTGAAGAGACA---GAGGAGATGCGAGGATTAACCTCG 125
QY 1257 GGAACAACACAGGATTTTTCCTGCCCCCTGCTCCAGTNCAGTNGGCTGGCCGACCCGC 1316
Db 126 GAAACAAACAGGATTTCTCTGCCCCCTGCTCCAG-----TCGAGTTGGCCTG-ACCCGC 177
QY 1317 TTGANTCAGTGACCAATTTGTGGCAGANCAAGGGAGAGCAGCTTCCAGCCTGGGTGAGA 1376
Db 178 TTGA-TCAGTGACCAATTTGTGGCAGA-CAGGGAGAGCAGCTTCCAGCCTGGGTGAGA 235
QY 1377 AGGGTGGGCGAGCCCTTGGCCCCCTCACCTNCCAGGCTGCTGTGNNAGAGTGTCAAGTG 1436
Db 236 AGGGTGGGCGAGCCCTTGGCCCCCTCACCT-CCAGGCTGCTGTG-AGAGTGTCAAGTG 293
QY 1437 TGTAAGGNCACCAANTCAGGNTTCAGTGACAGAACAGGTCAGGTCATGCCGCC 1496
Db 294 TGTAAAGGCCCAAACTC---AGGTTCAAGTCAGACAAACAGGT-CAGCAGTATGCCGCC 349
QY 1497 GNTAGTTAAANNNGGGGCCCTCTNAAACCCCTTGCTNNGCCTNCACTNCGCCAGCTCA 1556
Db 350 GTAAGTTAA---GGGGCCCTCTAAACCCCTTGCTGCCCTCA---CCTGGCCAGCTCA 402
QY 1557 NCCCTTTTGGGTGTAGGGGAAAAGATGCTGACCCCTGGGAAGGCTWCCCTGGTAGAAT 1616
Db 403 -CCCTTTTGGGTGTAGGGGAAAAGATGCTGACCCCTGGGAAGGCT-CCCTGGGTAGAA 460
QY 1617 ACACACACTTTTCAGGTTGTGCAACACAGGTCCTGAGTTGACCTCTGTTACGCCAAG 1676
Db 461 ACACACACTTTTCAGGTTGTGCAACACAGGTCCTGAGTTGACCTCTGTTACGCCAAG 520
QY 1677 GACCAAGAAAGTGTGTAAGTGAAGTGTCTCAGTNCCTCCAGACATGTGCCCTTTGCT 1736
Db 521 GACCAAGAAAGTGTGTAAGTGAAGTGTCTCAGT-CCCCAGACATGTGCCCTTTGCT 579
QY 1737 GCTGCTACCACTTTTCCCCAGAGCAGGCCCCGAGCCCTTTCAGGCCGAGCAGCTGCC 1796
Db 580 GCTGCTACCACTTTTCCCCAGAGCAGGCCCCGAGCCCTTTCAGGCCGAGCAGCTGCC 639
QY 1797 CCAGACTCGTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGATGAGGATATGTC 1856
Db 640 CCAGACTCGTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGATGAGGATATGTC 699
QY 1857 CTGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAG 1916
Db 700 CTGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAA--GGAAAGCAGCAGAGAG 757
QY 1917 AGGYTCCGGGCCCCAGNCCCACTNATCAGTGTNCCAGGCTGTGTTNCCCGAGNAG 1976

[illegible]

```

RESULT 9
US-10-087-192-884
; Sequence 884, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: MORRIS, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 884
; LENGTH: 2828
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-884

```

Query Match 25.6%; Score 672.8; DB 13; Length 2828;

Best Local Similarity 81.8%; Pred. No. 1.6e-196; Matches 798; Conservative 0; Mismatches 157; Indels 21; Gaps 1;	
QY	1 CCCCTCTCACAGCCCAAGGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
DB	1873 CACCCCTCACAGCCCAAGGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCGTCCACCGAGCAT 1932
QY	61 CTGAGCGGAGCTGGGAGGGAAGGTGAAACGGGCACTACAGCAAGTGGGAGGTCTGGAAGA 120
DB	1933 CTGCAATGGAGCTTCGGAGGAAATGGGCAAGCACTACAGGAAAGTGGGAGGTCTGAAA 1992
QY	121 GCCCTTGGAGGGAGAGATATAAAGAACCAAGACATCCACGGCCAATCAAGCCAATTACC 180
DB	1993 GCCCTTGGAAAGAGAGATATAAAGAACCAAGACCTCCACCCCAAGACCAAGCCACTGCC 2052
QY	181 ACCAGACCTCTCCATGGCCAGCCAGAGAGCTTTGCGCAAGGGGCCCAAGGCCCCCGGCCAC 240
DB	2053 ACCAGACCTCTACTCTCCGCCGAGAGA-----GNAACCCACCAG 2091
QY	241 CTGAGGAGACAAAGGGAGAGCCCTTAAGTCTCAGCTCTCTCTCCACAGAGCCGCCAG 300
DB	2092 CCAAGGCCAACACAGACGGGGCTCTGAGCCTCAGCTCTCTACCGCCAGAACCAACCCAG 2151
QY	301 AGCCAACAGTCTCTCTCTTGACTTTTCGCAAGAGGAGGTCTGGGATGTGGGACCTT 360
DB	2152 AACCGAGCAAGCGCCAGCCCTGAACTCTGAGCAAGGAGGAGTCTGGCAATGGGAAACCCC 2211
QY	361 TACTCTGTCTCCTCTGGAGCCAGCCCTTGCCAGAAACCCCAAGCTCACAGAGCGGAAAG 420
DB	2212 TGCTCTGTCTCTCTCTGGACCCAGCACTTGCCAAAGGCCCCAGCTTCCAGACCGGAGG 2271
QY	421 CAACCGTCCGGAGCAGGAACTGACGAGCTGGGAAATGAAATTAATTCCTCAACAGCTGT 480
DB	2272 CAACCTTTGCCAGAGCTGGAGCTACAGCAACTGGAGATAGAACTGTTTCTCAACAGCTGT 2331
QY	481 CCCAGCCATTTCTCTGGAGGAGCAGAGCAATTTCTCTGTCGCTCAGCATTCGACAGCC 540
DB	2332 CCCAGCGTGTCTCTGGAGGAAACAGAAACAATCTCTCTCTGCTCAGCATTCGACAGCC 2391
QY	541 TCTCCTCTCGATGACAGTGAGAGAAACCCATCAAAAGCCCTCTCAAAGCTCGCGGACA 600
DB	2392 TCTCGTGTCTAGATGACAGTGAGAAAGAAATCCATCGAAGGCCCTCAGAGCTCACGGACA 2451
QY	601 CCTTGAGCTCAGGGGTACACTCTCTGAGCAGCCAGGCTCGAAGCTCAGCTGGA 660
DB	2452 CCTTGAGTTCTGCGGTGCATCTTTGGAAACAGCCAAGCTCAGGCAAGAACTCAGCTGCA 2511
QY	661 ACATGGTGTGCGCCGGGGGGCGCCACCGACACCCCAAGCTATTCTCAATGGGTGGAAG 720
DB	2512 GCACGGCGCTGGCGCCGGGGGGCGGCTTACTGACATCCCGAGCTACTTCAACGGGGTCAAG 2571
QY	721 TCCAAATACAGTCTCTTAATTGGTGAACACTCTGCACATCCGGGAGTTCCACCGGGTCAAAG 780
DB	2572 TCCAGATCCAGTCTCTCAATGGGACACCTGCATATCCGGGAATTCACACCGGTCAAGG 2631
QY	781 TGGGAGACATCGCCACTGGGCACTGAGCGAAGATCCCAAGCTGAGCTTTCAGCTTGGTCA 840
DB	2632 TGGGAGACATTGCAACGGGCATCAGCAGCCAGATCCCAAGCCACAGCTTTTCAGCTTGGTGA 2691
QY	841 CCAAGACGGGAGCCTGTTTCGCTACGACATGGAGGTGCCAGCTCGGGCATTCGACCTGC 900
DB	2692 CCAAGATGGACAGCCTGTTTGTATGACATGGAGGTGCCAGCTCGGGCATTCGACCTGC 2751
QY	901 AGTGCACTTGGCCCTTGATGGCAGCTTCCCTCTGGAGCTTGGAGGGTCAAGCATTGGCCAGC 960
DB	2752 AGTGCACTTGGCCCTTGATGGCAGCTTTCCTTGGACCTGGAGGGTCAAGCATTGTTCAGC 2811
QY	961 TGGAGAACAGGCCCTA 976
DB	2812 TGGAGAACAGGCCCTA 2827

RESULT 10

```
US-10-108-260A-1754/c
; Sequence 1754, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1754
; LENGTH: 2475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1754

Query Match      13.3%; Score 349.4; DB 17; Length 2475;
Best Local Similarity 99.7%; Pred. No. 1.2e-96;
Matches 350; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 102 CAAAGTGGGAGGTCTGAAGAGCCCTTGGAGGGGAGAGTATAAAGAACCAAGACATCCACCG 161
DB 2179 CTAGTGGGAGGTCTGAGAGCCCTTGGAGGGGAGAGTATAAAGAACCAAGACATCCACCG 2120

QY 162 CCAATCAAGCCAAATTACACAGACCTTCATGCCCGCCGAGAGAGCTTTCGCCAAGG 221
DB 2119 CCAATCAAGCCAAATTACACAGACCTTCATGCCCGCCGAGAGAGCTTTCGCCAAGG 2060

QY 222 GCCCAGGCGCCGCCAGCTCAGAGACCAAGCAGAGCCCTTAAGCTCCAGCCTCCT 281
DB 2059 GCCCAGGCGCCGCCAGCTCAGAGACCAAGCAGAGCCCTTAAGCTCCAGCCTCCT 2000

QY 282 CTCACACAGAGCCCGCAGAGCAACAAAGTCTCTCTCTTGCATTTGAGCAAGGAGGAG 341
DB 1999 CTCACACAGAGCCCGCAGAGCAACAAAGTCTCTCTCTTGCATTTGAGCAAGGAGGAG 1940

QY 342 TCTGGATGTGGAAACCTTACTCTGTCTCTCTGGAGCCAGCCCTGCCAGAAACCCC 401
DB 1939 TCTGGATGTGGAAACCTTACTCTGTCTCTCTGGAGCCAGCCCTGCCAGAAACCCC 1880

QY 402 AGCTACACAGAGCGGAAAGCAACCGTCCGAGCAGAGCAACTGCAGAGCTG 452
DB 1879 AGCTACACAGAGCGGAAAGCAACCGTCCGAGCAGAGCAACTGCAGAGCTG 1829

RESULT 11
US-09-918-995-2585
; Sequence 2585, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2585
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(499)
; OTHER INFORMATION: n = A, T, C or G
US-09-918-995-2585

Query Match      9.2%; Score 242.2; DB 10; Length 499;
Best Local Similarity 90.9%; Pred. No. 8.3e-64;

US-10-108-260A-1754/c
; Sequence 1754, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1754
; LENGTH: 2475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1754

Query Match      13.3%; Score 349.4; DB 17; Length 2475;
Best Local Similarity 99.7%; Pred. No. 1.2e-96;
Matches 350; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 102 CAAAGTGGGAGGTCTGAAGAGCCCTTGGAGGGGAGAGTATAAAGAACCAAGACATCCACCG 161
DB 2179 CTAGTGGGAGGTCTGAGAGCCCTTGGAGGGGAGAGTATAAAGAACCAAGACATCCACCG 2120

QY 162 CCAATCAAGCCAAATTACACAGACCTTCATGCCCGCCGAGAGAGCTTTCGCCAAGG 221
DB 2119 CCAATCAAGCCAAATTACACAGACCTTCATGCCCGCCGAGAGAGCTTTCGCCAAGG 2060

QY 222 GCCCAGGCGCCGCCAGCTCAGAGACCAAGCAGAGCCCTTAAGCTCCAGCCTCCT 281
DB 2059 GCCCAGGCGCCGCCAGCTCAGAGACCAAGCAGAGCCCTTAAGCTCCAGCCTCCT 2000

QY 282 CTCACACAGAGCCCGCAGAGCAACAAAGTCTCTCTCTTGCATTTGAGCAAGGAGGAG 341
DB 1999 CTCACACAGAGCCCGCAGAGCAACAAAGTCTCTCTCTTGCATTTGAGCAAGGAGGAG 1940

QY 342 TCTGGATGTGGAAACCTTACTCTGTCTCTCTGGAGCCAGCCCTGCCAGAAACCCC 401
DB 1939 TCTGGATGTGGAAACCTTACTCTGTCTCTCTGGAGCCAGCCCTGCCAGAAACCCC 1880

QY 402 AGCTACACAGAGCGGAAAGCAACCGTCCGAGCAGAGCAACTGCAGAGCTG 452
DB 1879 AGCTACACAGAGCGGAAAGCAACCGTCCGAGCAGAGCAACTGCAGAGCTG 1829

RESULT 12
US-09-783-590-11852
; Sequence 11852, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11852
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (5)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (7)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (28)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (61)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (95)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (100)
; OTHER INFORMATION: n equals a,t,g, or c

Matches 288; Conservative 1; Mismatches 23; Indels 5; Gaps 3;

QY 811 AGATCCAGCTGCAGCCTTCAGCTTGGTTCACCAAGACGGGAGAGCTGTCGTACGACA 870
DB 182 AGATCCAGCTGCAGCCTTCAGCTTGGTTCACCAAGACGGGAGAGCTGTCGTACGACA 241

QY 871 TGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACACTGGGCCCTGATGGCAGCTTCG 930
DB 242 TGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACACTGGGCCCTGATGGCAGCTTCG 301

QY 931 CCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTAACCCCTGCCCTCCAC 990
DB 302 CCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTAACCCCTGCCCTCCAC 361

QY 991 CGCCGGCTCCACACTCGCGGAAGAGAGCCTTCTCTGTCGTCACGATGCTGCTCCCTGAAA 1050
DB 362 CGCCGGCTCCACACTCGCGGAAGAGAGCCTTCTCTGTCGTCACGATGCTGCTCCCTG-AA 419

QY 1051 ACACAGGCTCAGCCTTCCCGAGGGGATYTGNCAGCCCCCGGCTCARCAGNTGGGAACC 1110
DB 420 ACACAGGCTCAGCCTTCCCGATGG--ATCTGCGAGCCCCCGGCTCAGCAGAGGACCCAG 476

QY 1111 AGGCGCTCGNCAGCAG 1127
DB 477 GGCCTCGCAGCAGCAAG 493
```

```
; NAME/KEY: misc feature
; LOCATION: (130)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (195)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (228)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (247)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (250)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (268)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (277)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (287)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (293)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-11852

Query Match      7.3%; Score 193; DB 9; Length 296;
Best Local Similarity 88.2%; Pred. No. 1.1e-48;
Matches 224; Conservative 0; Mismatches 28; Indels 2; Gaps 2;

Qy 1558 CCCCTTTTGGGTGTAGGGGAAAGAAATGCTGACCCCTGGGAAAGCTWCCTGGTAGAATA 1617
Db 43 CCCCTTTTGGGTGTAGGGGAAAGAAAGAAATGCTGACCCCTGGGAAAGCTWCCTGGTAGAATA 102

Qy 1618 CACCACACTTTTTCAGGTTGTTGCAACACAGCTCCTGAGTTGACCTCTGTTTCAGCCAAAG 1677
Db 103 CACCACACTTTTTCAGGTTGTTGCAACANAGCTCCTGAGTTGACCTCTGTTTCAGCCAAAG 162

Qy 1678 ACCAAGAGTGTTGTAAGTGAAGTGTTCAGTNCCTGAGTCCCGAGATGCCCCCTTTG-CT 1736
Db 163 ACCAAGAGTGTTGTAAGTGAAGTGTTCAGTNCCTGAGTCCCGAGATGCCCCCTTTGCT 222

Qy 1737 GCTGCTACCACTCTTCCCGAGAGCAGCAGCCCGAGCCCC-TTTCAGGCCAGCACTGC 1795
Db 223 GCTGNTACCACTCTTCCCGAGAGNAGNAGGCCCGAGCCCCCTTTCAGGCCAGCACTGC 282

Qy 1796 CCCAGACTCGCTGG 1809
Db 283 CCCANATTCGNTGG 296

RESULT 13
US-10-087-192-883
; Sequence 883, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 883

; LENGTH: 42566
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(42566)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-883

Query Match      7.2%; Score 188.8; DB 13; Length 42566;
Best Local Similarity 73.9%; Pred. No. 1.4e-46;
Matches 263; Conservative 0; Mismatches 72; Indels 21; Gaps 1;

Qy 104 AGTGGAGGTTCTGAAGAGCCCTTGGAGGGGAGATATAAAGAACCAAGACATCCACCCGC 163
Db 27368 AGTGGAGGTTCTGAAGAGCCCTTGGAAAGAGGAGATATAAAGAACCAAGACATCCACCCGC 27427

Qy 164 AATCAAGCCCAATTACCAACAGACAGCCCTCCATGCCCAGCCGAGAGAGCTTTCCGCCAAGGC 223
Db 27428 AGACCAAGCCCACTGCCCAGACAGCCCTTACTCTCCGCGAGAGA----- 27472

Qy 224 CCCAGGCCCGCGGCGAGCTGAGGAGACAAAGAGCCAGAGCCCTTCTCTTCTTGAAGAGGAGGAGTC 283
Db 27473 -----GAACCCAGCCAGCCAGGCGCAAGCCGAGCCCTTGAACCTGAGCAAGGAGGAGTC 27526

Qy 284 CCCACAGAGCCCCAGAGCCAAACAAAGTCTCTCTCTTGAAGAGGAGGAGTC 343
Db 27527 ACCGCCAGAACCCAGCAAGCCGAGCAAGCCGAGCCCTTGAACCTGAGCAAGGAGGAGTC 27586

Qy 344 TGGATGTGGAAACCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 403
Db 27587 TGGCACAATGGAAACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 27646

Qy 404 CTACACAGAGCGGAAAGCAACCGCTCCCGAGAGCAAGTGCAGCAGCTGGAAATAG 459
Db 27647 CTTCACAGAGCGGAGGCAACCTTGCAGAGCTGGAGCTGAGCAACTGGAGATAG 27702

RESULT 14
US-10-357-930-22052/c
; Sequence 22052, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22052
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-22052
```

Mon Jun 13 13:27:12 2005

Query Match 3.4%; Score 88.4; DB 20; Length 455;
Best Local Similarity 98.9%; Pred. No. 2.6e-16;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 458 AGAATTATTCCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGGAGGAGCAAAATTCCT 517
Db |||||||
227 AGAATTATTCCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGGAGGAGCAAAATTCCT 168
Db |||||||

QY 518 CTCGTGCCTCAGCATCGACGCTCTCCCT 547
Db |||||||
167 CTCGTGCCTCAGCATCGACGCTCTCCCT 138
Db |||||||

RESULT 15
US-10-357-930-27909/c
; Sequence 27909, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27909
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-27909

Query Match 3.4%; Score 88.4; DB 20; Length 455;
Best Local Similarity 98.9%; Pred. No. 2.6e-16;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 458 AGAATTATTCCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGGAGGAGCAAAATTCCT 517
Db |||||||
227 AGAATTATTCCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGGAGGAGCAAAATTCCT 168
Db |||||||

QY 518 CTCGTGCCTCAGCATCGACGCTCTCCCT 547
Db |||||||
167 CTCGTGCCTCAGCATCGACGCTCTCCCT 138
Db |||||||

Search completed: June 11, 2005, 06:13:06
Job time : 1014.86 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 06:13:19 ; Search time 6580.54 Seconds
(without alignments)
16321.674 Million cell updates/sec

Title: US-09-155-676B-3
Perfect score: 2631
Sequence: 1 ccccttcacagccagccg.....accgtcgacctcgagggggg 2631

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1:	/cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq:*
2:	/cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq:*
3:	/cgn2_6/ptodata/1/pna/PCTUS3_COMB.seq:*
4:	/cgn2_6/ptodata/1/pna/US06_COMB.seq:*
5:	/cgn2_6/ptodata/1/pna/US07_COMB.seq:*
6:	/cgn2_6/ptodata/1/pna/US080_COMB.seq:*
7:	/cgn2_6/ptodata/1/pna/US081_COMB.seq:*
8:	/cgn2_6/ptodata/1/pna/US082_COMB.seq:*
9:	/cgn2_6/ptodata/1/pna/US083_COMB.seq:*
10:	/cgn2_6/ptodata/1/pna/US084_COMB.seq:*
11:	/cgn2_6/ptodata/1/pna/US085_COMB.seq:*
12:	/cgn2_6/ptodata/1/pna/US086_COMB.seq:*
13:	/cgn2_6/ptodata/1/pna/US087_COMB.seq:*
14:	/cgn2_6/ptodata/1/pna/US088_COMB.seq:*
15:	/cgn2_6/ptodata/1/pna/US089_COMB.seq:*
16:	/cgn2_6/ptodata/1/pna/US090_COMB.seq:*
17:	/cgn2_6/ptodata/1/pna/US091_COMB.seq:*
18:	/cgn2_6/ptodata/1/pna/US092A_COMB.seq:*
19:	/cgn2_6/ptodata/1/pna/US092B_COMB.seq:*
20:	/cgn2_6/ptodata/1/pna/US093A_COMB.seq:*
21:	/cgn2_6/ptodata/1/pna/US093B_COMB.seq:*
22:	/cgn2_6/ptodata/1/pna/US094_COMB.seq:*
23:	/cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
24:	/cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
25:	/cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
26:	/cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
27:	/cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
28:	/cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
29:	/cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
30:	/cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
31:	/cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
32:	/cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
33:	/cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
34:	/cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
35:	/cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
36:	/cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
37:	/cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
38:	/cgn2_6/ptodata/1/pna/US098D_COMB.seq:*
39:	/cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
40:	/cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
41:	/cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
42:	/cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
43:	/cgn2_6/ptodata/1/pna/US099E_COMB.seq:*

44:	/cgn2_6/ptodata/1/pna/US099F_COMB.seq:*
45:	/cgn2_6/ptodata/1/pna/US099G_COMB.seq:*
46:	/cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
47:	/cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
48:	/cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
49:	/cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
50:	/cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
51:	/cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
52:	/cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
53:	/cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
54:	/cgn2_6/ptodata/1/pna/US104A_COMB.seq:*
55:	/cgn2_6/ptodata/1/pna/US104B_COMB.seq:*
56:	/cgn2_6/ptodata/1/pna/US105A_COMB.seq:*
57:	/cgn2_6/ptodata/1/pna/US105B_COMB.seq:*
58:	/cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
59:	/cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
60:	/cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
61:	/cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
62:	/cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
63:	/cgn2_6/ptodata/1/pna/US108A_COMB.seq:*
64:	/cgn2_6/ptodata/1/pna/US108B_COMB.seq:*
65:	/cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
66:	/cgn2_6/ptodata/1/pna/US109B_COMB.seq:*
67:	/cgn2_6/ptodata/1/pna/US109C_COMB.seq:*
68:	/cgn2_6/ptodata/1/pna/US110_COMB.seq:*
69:	/cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
70:	/cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
71:	/cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
72:	/cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
73:	/cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
74:	/cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
75:	/cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
76:	/cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
77:	/cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
78:	/cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
79:	/cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
80:	/cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
81:	/cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
82:	/cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
83:	/cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
84:	/cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
85:	/cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
86:	/cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
87:	/cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
88:	/cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
89:	/cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
90:	/cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
91:	/cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
92:	/cgn2_6/ptodata/1/pna/US6023A_COMB.seq:*
93:	/cgn2_6/ptodata/1/pna/US6023B_COMB.seq:*
94:	/cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
95:	/cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
96:	/cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
97:	/cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
98:	/cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
99:	/cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
100:	/cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
101:	/cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
102:	/cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
103:	/cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
104:	/cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
105:	/cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
106:	/cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
107:	/cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
108:	/cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
109:	/cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
110:	/cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
111:	/cgn2_6/ptodata/1/pna/US6041_COMB.seq:*
112:	/cgn2_6/ptodata/1/pna/US6042_COMB.seq:*
113:	/cgn2_6/ptodata/1/pna/US6043_COMB.seq:*
114:	/cgn2_6/ptodata/1/pna/US6044_COMB.seq:*
115:	/cgn2_6/ptodata/1/pna/US6045_COMB.seq:*
116:	/cgn2_6/ptodata/1/pna/US6046_COMB.seq:*

QY 241 CTGAGGAGCAACAGGCGAGAGCCCTTAAGCTCCAGCTCTCTCTCCACCAGAGCCCCAG 300
Db 241 CTGAGGAGCAACAGGCGAGAGCCCTTAAGCTCCAGCTCTCTCTCCACCAGAGCCCCAG 300
QY 301 AGCCAAACAAGTCTCTCCCTTGAATTTGACAAAGAGGAGTCTGGGATGAGGAAACCT 360
Db 301 AGCCAAACAAGTCTCTCCCTTGAATTTGACAAAGAGGAGTCTGGGATGAGGAAACCT 360
QY 361 TACTCTGTCTCTCTGGAGCAGCCCTTGCCAGAAACCCAGCTCCAGAGCGGAAAG 420
Db 361 TACTCTGTCTCTCTGGAGCAGCCCTTGCCAGAAACCCAGCTCCAGAGCGGAAAG 420
QY 421 CAACGGTCCCGAGCAGGAACTGACAGAGCTGGAAATAGAATTAATCTCAACAGCTGT 480
Db 421 CAACGGTCCCGAGCAGGAACTGACAGAGCTGGAAATAGAATTAATCTCAACAGCTGT 480
QY 481 CCCAGCAATTTTCTCTGGAGCAGGAGCAAAATTTCTCGTGCTCAGCATGCACAGCC 540
Db 481 CCCAGCAATTTTCTCTGGAGCAGGAGCAAAATTTCTCGTGCTCAGCATGCACAGCC 540
QY 541 TCTCCCTGTGCGATGACAGTGAGAGAAACCCATCAAAAGGCTCTCAAAAGCTTCGGGACA 600
Db 541 TCTCCCTGTGCGATGACAGTGAGAGAAACCCATCAAAAGGCTCTCAAAAGCTTCGGGACA 600
QY 601 CCCTGAGCTCAGGCGTACATCTCTGGAGCAGCCAGGCGGAGCTCGAAGCTCCAGCTGGA 660
Db 601 CCCTGAGCTCAGGCGTACATCTCTGGAGCAGCCAGGCGGAGCTCGAAGCTCCAGCTGGA 660
QY 661 ACATGGTGTGCGCCCGGGGGCGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 720
Db 661 ACATGGTGTGCGCCCGGGGGCGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 720
QY 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCGGGAGTTCACCGGTCAAAG 780
Db 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCGGGAGTTCACCGGTCAAAG 780
QY 781 TGGAGACATCGCACTGGCATAGCAGCAGATCCAGCTGCAGCTTCAGCTTGTGCA 840
Db 781 TGGAGACATCGCACTGGCATAGCAGCAGATCCAGCTGCAGCTTCAGCTTGTGCA 840
QY 841 CCAAGACGGCAGCCTTTCGCTACGACATGGAGGTGCCAGATCGGGCATCGACCTGC 900
Db 841 CCAAGACGGCAGCCTTTCGCTACGACATGGAGGTGCCAGATCGGGCATCGACCTGC 900
QY 901 AGTGCACTGGCCCTGTAGGCGAGTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 960
Db 901 AGTGCACTGGCCCTGTAGGCGAGTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 960
QY 961 TGGAGAACAGCCCTAACCCCTGCCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT 1020
Db 961 TGGAGAACAGCCCTAACCCCTGCCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT 1020
QY 1021 TCCTGTCTGGTGACAGATGCTGCCCTGAAACACAGCTCAGCGTTCACAGGGGATYTG 1080
Db 1021 TCCTGTCTGGTGACAGATGCTGCCCTGAAACACAGCTCAGCGTTCACAGGGGATYTG 1080
QY 1081 NCCAGCCCCCGGCTCARGNTGGGAACAGGSCCTCGNCAGCAGCNAAGTNGGGG 1140
Db 1081 NCCAGCCCCCGGCTCARGNTGGGAACAGGSCCTCGNCAGCAGCNAAGTNGGGG 1140
QY 1141 CAAGCNAAGATGCTCCAGGATTTACANCTGAGCCNTGCCCANCCCTGCTGAADA 1200
Db 1141 CAAGCNAAGATGCTCCAGGATTTACANCTGAGCCNTGCCCANCCCTGCTGAADA 1200
QY 1201 AAACAYTNCGCCACGTGAAGAGACAGAGAGATGNCAGGAGTTNACCTYGGGAA 1260
Db 1201 AAACAYTNCGCCACGTGAAGAGACAGAGAGATGNCAGGAGTTNACCTYGGGAA 1260
QY 1261 ACAAACAGGGATCTTTNTTCTGCCCTGTCCAGTNCAGTGGCCCTGNACCCGCTTG 1320
Db 1261 ACAAACAGGGATCTTTNTTCTGCCCTGTCCAGTNCAGTGGCCCTGNACCCGCTTG 1320
QY 1321 ANTCAGTGACCATTTGTTGGCAGANCAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGAGG 1380

Db 1321 ANTCAGTGACCATTTGTTGGCAGANCAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGGG 1380
QY 1381 GTGGCGAGCCCTTCGGCCCTCACCTTNCAGGCTGCTGTGNAGAGTGTCAAGTGTGTA 1440
Db 1381 GTGGCGAGCCCTTCGGCCCTCACCTTNCAGGCTGCTGTGNAGAGTGTCAAGTGTGTA 1440
QY 1441 AGGNCCTCAANCTCAGGNTTCAAGTGAGAACACAGGNTCAGCAGGTATGCCCGCCGNTA 1500
Db 1441 AGGNCCTCAANCTCAGGNTTCAAGTGAGAACACAGGNTCAGCAGGTATGCCCGCCGNTA 1500
QY 1501 GGTAAANNNGGGCCCTCTNAAACCCCTTGCTTNGGCTTNCACCTNGGCCAGCTCANCC 1560
Db 1501 GGTAAANNNGGGCCCTCTNAAACCCCTTGCTTNGGCTTNCACCTNGGCCAGCTCANCC 1560
QY 1561 CTTTGTGGTGTAGGGGAAAAGATGCTGACCCCTGGGAAGGCTWCCCTGGTAGAATACAC 1620
Db 1561 CTTTGTGGTGTAGGGGAAAAGATGCTGACCCCTGGGAAGGCTWCCCTGGTAGAATACAC 1620
QY 1621 CACATTTTTCAGGTTGTGCAACACAGGTCTGAGTTGACCTCTGGTTCAGCCAGGACC 1680
Db 1621 CACATTTTTCAGGTTGTGCAACACAGGTCTGAGTTGACCTCTGGTTCAGCCAGGACC 1680
QY 1681 AAAGAGGTGTAAAGTGAAGTGGTCTCAGTNCCTCCAGACATGTGCCCTTTGTGTGTG 1740
Db 1681 AAAGAGGTGTAAAGTGAAGTGGTCTCAGTNCCTCCAGACATGTGCCCTTTGTGTGTG 1740
QY 1741 GCTACCACTCTTCCCCAGAGCAGGCCCCGAGCCCCCTTCAGGCCACGACTGCCCCAG 1800
Db 1741 GCTACCACTCTTCCCCAGAGCAGGCCCCGAGCCCCCTTCAGGCCACGACTGCCCCAG 1800
QY 1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGATGAGGATATGCTCTGA 1860
Db 1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGATGAGGATATGCTCTGA 1860
QY 1861 CAGGAAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGAGCY 1920
Db 1861 CAGGAAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGAGCY 1920
QY 1921 TCCGGCCCCCAGNCCCCTACTNATCAGTGTNCCAGCGTCTNGGTTNCCCAGNAGACA 1980
Db 1921 TCCGGCCCCCAGNCCCCTACTNATCAGTGTNCCAGCGTCTNGGTTNCCCAGNAGACA 1980
QY 1981 GCTNCAGNCATCANCACTGACATNCACTTNGCCCTGCCCTNGGCCANGAGGTTACTG 2040
Db 1981 GCTNCAGNCATCANCACTGACATNCACTTNGCCCTGCCCTNGGCCANGAGGTTACTG 2040
QY 2041 CCGNACGGCACTTTGCACTNTCTGATGNACCTCAAAGCACTTTTCAATGGCTNGCCCTCT 2100
Db 2041 CCGNACGGCACTTTGCACTNTCTGATGNACCTCAAAGCACTTTTCAATGGCTNGCCCTCT 2100
QY 2101 GCAGGNCAGGNCAGGNCAGTGACATCTGTAGNAGCATANGCAAGCCAGAGATGG 2160
Db 2101 GCAGGNCAGGNCAGGNCAGTGACATCTGTAGNAGCATANGCAAGCCAGAGATGG 2160
QY 2161 GGTGNAAGGNCACAGTCTTGAGCTGTCCANCATGATGACTNCTCAAACCTCTTN 2220
Db 2161 GGTGNAAGGNCACAGTCTTGAGCTGTCCANCATGATGACTNCTCAAACCTCTTN 2220
QY 2221 NCCAGNATTTCTTAAGAATPAGCANCCCCCTTNNCCCCATTTGCCCAGCTTAGCCTTCT 2280
Db 2221 NCCAGNATTTCTTAAGAATPAGCANCCCCCTTNNCCCCATTTGCCCAGCTTAGCCTTCT 2280
QY 2281 CCAGGGGAGCTANCTCAGGACTCACTGATCATTAATCAGCTGTGNAATCGTCAGGGG 2340
Db 2281 CCAGGGGAGCTANCTCAGGACTCACTGATCATTAATCAGCTGTGNAATCGTCAGGGG 2340
QY 2341 TGTCTGTAGCTCAACCTCTGGGCGAGGGAGCGCGAGACTCCGTTGGGAGAGCTCAT 2400
Db 2341 TGTCTGTAGCTCAACCTCTGGGCGAGGGAGCGCGAGACTCCGTTGGGAGAGCTCAT 2400
QY 2401 TCCCATCTTTCGCAAGACAGCCTTTNGTCCAGCTGTCCACATTTGAGTCAAGCTGCTCCC 2460

Db 2401 TCCACATCTTGTCCAAAGACAGCCTTTNGTCCAGTGTCTCCATTCAGTGCAGAGTCTCCC 2460
Qy 2461 GGGGAGAGAGCCCGGCCCCCAGACATAAAGAACTGCAGCTTGGTACTGCAGAGTCTG 2520
Db 2461 GGGGAGAGAGCCCGGCCCCCAGACATAAAGAACTGCAGCTTGGTACTGCAGAGTCTG 2520
Qy 2521 GGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTGGGGTGATGACAAAATGTTAAAAAAG 2580
Db 2521 GGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTGGGGTGATGACAAAATGTTAAAAAAG 2580
Qy 2581 GCCTTCGTGGCTCGAATCAAGCTTATCGATACCGTGCAGCTTCGAGGGGGG 2631
Db 2581 GCCTTCGTGGCTCGAATCAAGCTTATCGATACCGTGCAGCTTCGAGGGGGG 2631

RESULT 2

US-09-155-676A-3
; Sequence 3, Application US/09155676A
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/155,676A
; FILING DATE: 04-JAN-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IL97/00117
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117800
; FILING DATE: 02-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 119133
; FILING DATE: 26-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: WALLACH=21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2631 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-155-676A-3

Query Match 97.1%; Score 2553.8; DB 17; Length 2631;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCCTCTCACAGCCCGGCTCCTCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60

Db 1 CCCCTCTCACAGCCCGGCTCCTCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
Qy 61 CTGCAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGCTCTGAAGA 120
Db 61 CTGCAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGCTCTGAAGA 120
Qy 121 GCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCCAATCAAGCCAAATTACC 180
Db 121 GCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCCAATCAAGCCAAATTACC 180
Qy 181 ACCAGACCTTCATGCCAGCGGAGAGCTTTTCGCAAGGGGCCCAAGGGCCCGGGCCAG 240
Db 181 ACCAGACCTTCATGCCAGCGGAGAGCTTTTCGCAAGGGGCCCAAGGGCCCGGGCCAG 240
Qy 241 CTGAGGAGACAAAGGAGAGAGCCCTTAAGCTTCCAGCTTCTCCACAGAGCCCGCCAG 300
Db 241 CTGAGGAGACAAAGGAGAGAGCCCTTAAGCTTCCAGCTTCTCCACAGAGCCCGCCAG 300
Qy 301 AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCT 360
Db 301 AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCT 360
Qy 361 TACCTCTGTCTCTCTGGAGCGAGCCCTGCGAGAAACCCAGCTTCACAGAGCGGAAAG 420
Db 361 TACCTCTGTCTCTCTGGAGCGAGCCCTGCGAGAAACCCAGCTTCACAGAGCGGAAAG 420
Qy 421 CAACGCTCCCGAGCAGGAACTGCAGAGCTGGGAATAGAATTATTCCTCAACAGCCGTGT 480
Db 421 CAACGCTCCCGAGCAGGAACTGCAGAGCTGGGAATAGAATTATTCCTCAACAGCCGTGT 480
Qy 481 CCCAGCAATTTCTCTGGAGGAGAGGAGCAAAATTTCTCGTGCCTCAGCATTCGACAGCC 540
Db 481 CCCAGCAATTTCTCTGGAGGAGAGGAGCAAAATTTCTCGTGCCTCAGCATTCGACAGCC 540
Qy 541 TCTCCCTGTCCGATGACAGTGAAGAAGAACCCATCAAGGCTCTCAAGCTTCGCGGAGCA 600
Db 541 TCTCCCTGTCCGATGACAGTGAAGAAGAACCCATCAAGGCTCTCAAGCTTCGCGGAGCA 600
Qy 601 CCCTGAGCTCAGGGGTACACTCTTGAGAGCAGCGGCGAGGCTCGAAGCTCCAGCTGGA 660
Db 601 CCCTGAGCTCAGGGGTACACTCTTGAGAGCAGCGGCGAGGCTCGAAGCTCCAGCTGGA 660
Qy 661 ACATGGTCTGCGCGGGCGGCGCCACCCAGACACCCCAAGCTATTTCATGGTGTGAAG 720
Db 661 ACATGGTCTGCGCGGGCGGCGCCACCCAGACACCCCAAGCTATTTCATGGTGTGAAG 720
Qy 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCGGAGTTCCACCGGGTCAAAG 780
Db 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCGGAGTTCCACCGGGTCAAAG 780
Qy 781 TGGGAGACATCGCCACTGGCATCAGCAGCGAGATCCAGCTTCAGCTTGGTCA 840
Db 781 TGGGAGACATCGCCACTGGCATCAGCAGCGAGATCCAGCTTCAGCTTGGTCA 840
Qy 841 CCAAAGACGGCAGCCTCTGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC 900
Db 841 CCAAAGACGGCAGCCTCTGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC 900
Qy 901 AGTGACACTGGCCCCCTGATGGAGCTTCGCTGGAGTGGAGGTGAAGCAATGGCCAGC 960
Db 901 AGTGACACTGGCCCCCTGATGGAGCTTCGCTGGAGTGGAGGTGAAGCAATGGCCAGC 960
Qy 961 TGGGAAACAGGCCCTTAACCTTCCTCCACCGCGGCTCCACACTGCCGGAAGCAGCT 1020
Db 961 TGGGAAACAGGCCCTTAACCTTCCTCCACCGCGGCTCCACACTGCCGGAAGCAGCT 1020
Qy 1021 TCCTCTCGGTGCACGATGCTGCCCTGAAAACACAGGCTCAGCCCTTCCACGGGATYTG 1080
Db 1021 TCCTCTCGGTGCACGATGCTGCCCTGAAAACACAGGCTCAGCCGTTCACGGGATYTG 1080
Qy 1081 NCCAGCCCCCGGCTCCTACAGTGGGAACCCAGGGCTTCGNCAGCAGCAGGAGGAGG 1140

Db 1081 NCCAGCCCCCGGCTCARGNTGGGAACAGGGCCCTCGNCAGCNAGCNAAAGTNGGGG 1140
Qy 1141 CAAGCNAGAATGCTCCAGGATTTACANCTGAGCCCTGAGCCCTGCTGAADA 1200
Db 1141 CAAGCNAGAATGCTCCAGGATTTACANCTGAGCCCTGAGCCCTGCTGAADA 1200
Qy 1201 AAACAYTNCCGCCAGCTGAAGAGACAGAGAGAGATGNCAGGATTTNACCTTGGGAA 1260
Db 1201 AAACAYTNCCGCCAGCTGAAGAGACAGAGAGAGATGNCAGGATTTNACCTTGGGAA 1260
Qy 1261 ACAAACAGGATCTTTNTTCTGCCCCCTGCTCCAGTNGAGTGGCCCTGNAACCCGCTTGG 1320
Db 1261 ACAAACAGGATCTTTNTTCTGCCCCCTGCTCCAGTNGAGTGGCCCTGNAACCCGCTTGG 1320
Qy 1321 ANTCACTGACCATTTGTCAGANCAAGGAGAGAGAGCTTCCAGCCCTGGGTCAAGAGGG 1380
Db 1321 ANTCACTGACCATTTGTCAGANCAAGGAGAGAGAGCTTCCAGCCCTGGGTCAAGAGGG 1380
Qy 1381 GTGGCGAGCCCTTGGGCCCTTCAACCTTCCAGGTCGTGTGNAGAGTGTCAAGTGTGA 1440
Db 1381 GTGGCGAGCCCTTGGGCCCTTCAACCTTCCAGGTCGTGTGNAGAGTGTCAAGTGTGA 1440
Qy 1441 AGGNCNCAANCTCAGNTTTCAGTGCAAGACAGGTCNACAGGTATGCCCGCCGNTA 1500
Db 1441 AGGNCNCAANCTCAGNTTTCAGTGCAAGACAGGTCNACAGGTATGCCCGCCGNTA 1500
Qy 1501 GGTAAANNNGGGGGCCCTTNAACCCCTTGGCTTNGCCCTNACCTTNGCCAGCTCANCCC 1560
Db 1501 GGTAAANNNGGGGGCCCTTNAACCCCTTGGCTTNGCCCTNACCTTNGCCAGCTCANCCC 1560
Qy 1561 CTTTGGGTAGGGGAAAAGATGCTGACCCCTGGGAGAGCTTCCCTGGTAGAATACAC 1620
Db 1561 CTTTGGGTAGGGGAAAAGATGCTGACCCCTGGGAGAGCTTCCCTGGTAGAATACAC 1620
Qy 1621 CACACTTTTCAAGTGTGTCACACAGCTCTGAGTTGACCTCTGTTTCAGCCAGGACC 1680
Db 1621 CACACTTTTCAAGTGTGTCACACAGCTCTGAGTTGACCTCTGTTTCAGCCAGGACC 1680
Qy 1681 AAAGAAGTGTGAAGTGAAGTGTCTCAGTNCNCCAGACATGTGCCCTTTGCTGCTG 1740
Db 1681 AAAGAAGTGTGAAGTGAAGTGTCTCAGTNCNCCAGACATGTGCCCTTTGCTGCTG 1740
Qy 1741 GCTACCACTTCTCCAGAGCAGAGCCCGCCGAGCCCTTCCAGCCAGCAGCTGCCCCAG 1800
Db 1741 GCTACCACTTCTCCAGAGCAGAGCCCGCCGAGCCCTTCCAGCCAGCAGCTGCCCCAG 1800
Qy 1801 ACTGCTGGCACTCAGTTCCTCATCTGAAGGTGAAGGTGATGAGGATATGCTGA 1860
Db 1801 ACTGCTGGCACTCAGTTCCTCATCTGAAGGTGAAGGTGATGAGGATATGCTGA 1860
Qy 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAGNAAAGCAGCAGAGAGAGCY 1920
Db 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAGNAAAGCAGCAGAGAGAGCY 1920
Qy 1921 TCCGGCGCCCGAGNCCCACTNATCAGTGTNCCAGCGTGTGCTGCTGCTGCTGCTG 1980
Db 1921 TCCGGCGCCCGAGNCCCACTNATCAGTGTNCCAGCGTGTGCTGCTGCTGCTGCTG 1980
Qy 1981 GCTNAGNATCANCACTGACACTNCCCTTNGCCCTGCCCCCTGCCCCCTGCCCCCTG 2040
Db 1981 GCTNAGNATCANCACTGACACTNCCCTTNGCCCTGCCCCCTGCCCCCTGCCCCCTG 2040
Qy 2041 CCGNAGGCACTTGTGACNTGTGACNTGACCTTCAAGGACCTTTCATGCTGCTGCTGCTG 2100
Db 2041 CCGNAGGCACTTGTGACNTGTGACNTGACCTTCAAGGACCTTTCATGCTGCTGCTGCTG 2100
Qy 2101 GCAGGNCAGGNCAGGNCAGTGCACACTGTAGNAGCATANGCAAGCCAGGAGATGG 2160
Db 2101 GCAGGNCAGGNCAGGNCAGTGCACACTGTAGNAGCATANGCAAGCCAGGAGATGG 2160
Qy 2161 GGTNAAGGNCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
Db 2161 GGTNAAGGNCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220

Qy 2221 NCCAGNATTTCTTAAGATAGCANCCCCCTTNNCCCAATGCCCCAGCTTAGCCTTCT 2280
Db 2221 NCCAGNATTTCTTAAGATAGCANCCCCCTTNNCCCAATGCCCCAGCTTAGCCTTCT 2280
Qy 2281 CCCAGGGAGCTANTCTCAGGACTCAGCTAGCATTAATCAGCTGTGNAATGCTCAGGGG 2340
Db 2281 CCCAGGGAGCTANTCTCAGGACTCAGCTAGCATTAATCAGCTGTGNAATGCTCAGGGG 2340
Qy 2341 TGTCTGCTAGCTCAACCTCTCTGGGGCAGGGAGCCCGAGACTCCGTTGGGAGAGCTCAT 2400
Db 2341 TGTCTGCTAGCTCAACCTCTCTGGGGCAGGGAGCCCGAGACTCCGTTGGGAGAGCTCAT 2400
Qy 2401 TCCACATCTTGGCAAGACAGCTTTTNGTCCAGCTGTCCATTAATGAGTGCAGCTGCC 2460
Db 2401 TCCACATCTTGGCAAGACAGCTTTTNGTCCAGCTGTCCATTAATGAGTGCAGCTGCC 2460
Qy 2461 GGGGAGAGAGCCCGGCCCGCCAGCACAATAAGAACTGCAGCTTGGTACTGCAGAGTCTG 2520
Db 2461 GGGGAGAGAGCCCGGCCCGCCAGCACAATAAGAACTGCAGCTTGGTACTGCAGAGTCTG 2520
Qy 2521 GGTGTAGAGAACTCTTTGTAAGCAATAAGTTTGGGTGATGACAAATGTTAAAAAAG 2580
Db 2521 GGTGTAGAGAACTCTTTGTAAGCAATAAGTTTGGGTGATGACAAATGTTAAAAAAG 2580
Qy 2581 GCCTTCTGCTCGAATCAAGCTTTATCGATACCGCTCGACCTCGAGGGGG 2631
Db 2581 GCCTTCTGCTCGAATCAAGCTTTATCGATACCGCTCGACCTCGAGGGGG 2631

RESULT 3

US-09-155-676-6
; Sequence 6, Application US/09155676
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
; FACTOR (TRAF), THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/155,676
; FILING DATE: 04-JAN-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IL97/00117
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117800
; FILING DATE: 02-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 119133
; FILING DATE: 26-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: WALLACH-21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4596 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-155-676-6

Query Match 70.4%; Score 1851.2; DB 17; Length 4596;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2335; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

QY 1 CCCCTCTCACAGCCAGCCGATCCAAAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT 60
Db 2099 CCCCTCTCACAGCCAGCCGATCCAAAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT 2158

QY 61 CTGACGCGGAGCTGGAGGGAAGGTGAACCGGCACTACAGCAAGTGGGAGGTCTGAAGA 120
Db 2159 CTGACGCGGAGCTGGAGGGAAGGTGAACCGGCACTACAGCAAGTGGGAGGTCTGAAGA 2218

QY 121 GCCCTTGGAGGGAGATATAAGAAACCAAGACATCCACCGCCAATCAAGCCAAATTACC 180
Db 2219 GCCCTTGGAGGGAGATATAAGAAACCAAGACATCCACCGCCAATCAAGCCAAATTACC 2278

QY 181 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGGCCAG 240
Db 2279 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGGCCAG 2338

QY 241 CTGAGAGACAACAGCAGAGCCCTTAAGCTCCAGCTCCCTCCACAGAGCCCGCCAG 300
Db 2339 CTGAGAGACAACAGCAGAGCCCTTAAGCTCCAGCTCCCTCCACAGAGCCCGCCAG 2398

QY 301 AGCCAAACAAGTCTCCCTCCCTTACCTTTGAGCAAGAGGAGGTCTGGGATGTGGGAACCT 360
Db 2399 AGCCAAACAAGTCTCCCTCCCTTACCTTTGAGCAAGAGGAGGTCTGGGATGTGGGAACCT 2458

QY 361 TACCTCTGCTCCCTGGAGCCAGCCCTTGCCAGAAACCCAGCTCACCAGAGCCGGAAG 420
Db 2459 TACCTCTGCTCCCTGGAGCCAGCCCTTGCCAGAAACCCAGCTCACCAGAGCCGGAAG 2518

QY 421 CAACCTCCCGAGCAGGAGCTCAGCAGCTGGAATAGAAATTAATTCCTCAACAGCCTGT 480
Db 2519 CAACCTCCCGAGCAGGAGCTCAGCAGCTGGAATAGAAATTAATTCCTCAACAGCCTGT 2578

QY 481 CCCAGCCATTTCTCTGGAGGAGCAGGAGCAAAATCTCTGCTGCTCAGCATCGACGCC 540
Db 2579 CCCAGCCATTTCTCTGGAGGAGCAGGAGCAAAATCTCTGCTGCTCAGCATCGACGCC 2638

QY 541 TCTCCCTGTCGATGACAGTGAGAGAACCCATCAAGGCTCTCAAGCTCCGCGGACA 600
Db 2639 TCTCCCTGTCGATGACAGTGAGAGAACCCATCAAGGCTCTCAAGCTCCGCGGACA 2698

QY 601 CCCTGAGCTCAGGCTACACTCTCTGGAGCAGCAGGCGGAGGCTCGAAGCTCCAGTGA 660
Db 2699 CCCTGAGCTCAGGCTACACTCTCTGGAGCAGCAGGCGGAGGCTCGAAGCTCCAGTGA 2758

QY 661 ACATGGTGTGGCCGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 720
Db 2759 ACATGGTGTGGCCGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 2818

QY 721 TCCAAATACAGTCTCTTAATGGTGAACACTCGGAGTTCACCGGGTCAAG 780
Db 2819 TCCAAATACAGTCTCTTAATGGTGAACACTCGGAGTTCACCGGGTCAAG 2878

QY 781 TGGGAGACATCGCCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGTTGGTCA 840
Db 2879 TGGGAGACATCGCCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGTTGGTCA 2938

QY 841 CCAAGAGCGGCGAGCTGTTCGCTACGACATGAGGTTGCCAGCTCGGGCATCGACCTGC 900
Db 2939 CCAAGAGCGGCGAGCTGTTCGCTACGACATGAGGTTGCCAGCTCGGGCATCGACCTGC 2998

QY 901 AGTGACACTGGCCCTGATGAGCTTCGCTGGAGCTGGAGGCTCAAGCATGGCCAGC 960
Db 2999 AGTGACACTGGCCCTGATGAGCTTCGCTGGAGCTGGAGGCTCAAGCATGGCCAGC 3058

QY 961 TGGAGAACAGGCTTAACCTTCACCGCGGCTCCACATGCTCCCGGAAGAGCCT 1020
Db 3059 TGGAGAACAGGCTTAACCTTCACCGCGGCTCCACATGCTCCCGGAAGAGCCT 3118

QY 1021 TCTGCTCGGTGACCATGCTGCTCCCTGAAAACACACAGGCTCAGCCGTTCCAGGGGAT-- 1080
Db 3119 TCTGCTCGGTGACCATGCTGCTCCCTGAAAACACACAGGCTCAGCCGTTCCAGGGGAT-- 3176

QY 1081 NCCAGCCCCCGGCTCARCAGNTGGGAACACAGGCTTCGNCAGCNAAAGTNGGGG 1140
Db 3177 GCACAGCCCCCGGCTCA--CAGTGGGAACACAGGCTTCG--CAGCAGCAAGGTGGG 3229

QY 1141 CAAGCNAAGATGCTCCAGGATTTACANCTGAGCCCTGCCCANCCCTGCTCAADA 1200
Db 3230 CAAGCAGAAATGCTCCAGGATTTACACCTGAGCCCTGCCCCAC-----CCTGCTGA 3283

QY 1201 AAACAYTNCGCCACGCTGAAGAGACAGAGGAGATGNCAGAGTTNNACCTTYGGGAA 1260
Db 3284 AAACAYTNCGCCACGCTGAAGAGACAGAGGAGATGNCAGAGTTNNACCTTYGGGAA 3339

QY 1261 AAAAAACAGGATCTTTTCTGCTCCCTGCTCCAGTNCAGTGTGGCTGNAACCGCTTGG 1320
Db 3340 AAAAAACAGGATCTTTTCTGCTCCCTGCTCCAGT--CGAGTTGGCTGA--CCCGCTTG 3395

QY 1321 ANTCACTGACATTTGTTGGCAGANAGGGAGAGAGCTTCAGCCCTGGGTCAAGAGGG 1380
Db 3396 GATCAGTGACATTTGTTGGCAGA--CAGGGAGAGAGCTTCAGCCCTGGGTCAAGAGGG 3454

QY 1381 GTGGCGGAGCCCTTCGGCCCTCACCCTNCAGGCTCTGTGNAGAGTGTCAAGTGTGTA 1440
Db 3455 GTGGCGGAGCCCTTCGGCCCTCACCCT--CCAGGCTCTGTG--AGAGTGTCAAGTGTGTA 3512

QY 1441 AGGNCCCAAANCTCAGNTTTCAGTGCAAGAACAGGTNCAGAGGTATGCCGCCCGNTA 1500
Db 3513 AGGNCCCAAANCTC--AGGTTTCAGTGCAAGAACAGGT--CAGCAGGTATGCCGCCCGNTA 3568

QY 1501 GGTAAANNNGGGGCTCTTNAACCCCTTGCCTNGGCTTNCACCTNGCCAGCTCANCCC 1560
Db 3569 GTTAA---GGGGGCTCTTAAACCCCTTGCCTNGGCTTNCACCTNGCCAGCTCANCCC 3620

QY 1561 CTTTGGGTGTAGGGGAAAAGATGCTGAGCCCTGGGAAGGCTWCCCTGTGTGAGTAATAC 1620
Db 3621 CTTTGGGTGTAGGGGAAAAGATGCTGAGCCCTGGGAAGGCT--CCCTGTGTGAGTAATAC 3679

QY 1621 CACACTTTTCAGGTTGTGCAACACAGGTCTGAGTTGACCTGTGTTTCAGCCAGGACC 1680
Db 3680 CACACTTTTCAGGTTGTGCAACACAGGTCTGAGTTGACCTGTGTTTCAGCCAGGACC 3739

QY 1681 AAAGAGGTGTGAAGTGAAGTGTTCACGTNCCCGAGACATGTGCCCTTTTGTGCTG 1740
Db 3740 AAAGAGGTGTGAAGTGAAGTGTTCACGTNCCCGAGACATGTGCCCTTTTGTGCTG 3798

QY 1741 GCTACACTCTTCCCAGAGCAGGAGCCCGAGCCCTTTCAGGCCAGACTGCCCCAG 1800
Db 3799 GCTACACTCTTCCCAGAGCAGGAGCCCGAGCCCTTTCAGGCCAGACTGCCCCAG 3858

QY 1801 ACTCGCTGGCACTCAGTTTCCCTCATCTCTAAAGGTGAAGGTGATGACAGGATATGCTGA 1860
Db 3859 ACTCGCTGGCACTCAGTTTCCCTCATCTCTAAAGGTGAAGGTGATGACAGGATATGCTGA 3918

QY 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGAGCG 1920
Db 3919 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAA--GGAAAGCAGCAGAGAGAGCG- 3975

QY 1921 TCCGGCGCCCGAGNCCCACTNATCAGTGTNCCAGGCTGCTNGTTNCCCGAGNAGACA 1980
Db 3976 TCCGGCGCCCGAGNCCCACTNATCAGTGT---CCAGGCTGCTGTTTCCCGAGAGCAGC- 4031

QY 721 TCCAAATACAGTCTCTTAATGGTGAAACCTGCGACATCCGGGAGTTCCACGGGTCAAAG 780
Db 2819 TCCAAATACAGTCTCTTAATGGTGAAACCTGCGACATCCGGGAGTTCCACGGGTCAAAG 2878
QY 781 TGGGAGACATCGCCCATCTGGGATCAGCAGCAGATCCAGCTGCGAGCTTCAAGCTTGCTCA 840
Db 2879 TGGGAGACATCGCCCATCTGGGATCAGCAGCAGATCCAGCTGCGAGCTTCAAGCTTGCTCA 2938
QY 841 CCAAGAGCGGCGAGCTCTCGCTACGATCGAGGTGCCAGATCGGGGCATCCAGCTGC 900
Db 2939 CCAAGAGCGGCGAGCTCTCGCTACGATCGAGGTGCCAGATCGGGGCATCCAGCTGC 2998
QY 901 AGTGACACTGGCCCCCTGATGCGAGCTTCGCTGGAGCTGGAGGTCAAGCATGCCCCAGC 960
Db 2999 AGTGACACTGGCCCCCTGATGCGAGCTTCGCTGGAGCTGGAGGTCAAGCATGCCCCAGC 3058
QY 961 TGGAGAACAGGCGCTTAAACCTGCTCCATCCAGCGGCTCCACATGCGCGGAAAGAGGCT 1020
Db 3059 TGGAGAACAGGCGCTTAAACCTGCTCCATCCAGCGGCTCCACATGCGCGGAAAGAGGCT 3118
QY 1021 TCCTGCTCGGTGACAGATGCTGCTGAAACACAGAGGCTCAGCGCTTCCAGGGGATVTG 1080
Db 3119 TCCTGCTCGGTGACAGATGCTGCTGAAACACAGAGGCTCAGCGCTTCCAGGGGATVTG 3176
QY 1081 NCCAGCCCCCGGCTCARCAGNTGGGAACCCAGGGCTCCGNCAGCNAGNAAAGTNGGGG 1140
Db 3177 GCCAGCCCCCGGCTCA--CAGTGGGAACCCAGGGCTCG----CAGCAGCAAGGTGGGG 3229
QY 1141 CAAGNAGATGCTCCAGAGATTTCAACNCTGAGCCGCTGCGCCAGCCCTGCTGAADA 1200
Db 3230 GCAAGCAAGATGCTCCAGAGATTTCAACNCTGAGCCCTGCGCCAC-----CCTGCTGA 3283
QY 1201 AAACATYNNCCGACGTGAAGAGACAGAGGAGGATGNCAGGAGTTNNACTYGGGGA 1260
Db 3284 AAAAATATCCGCCACGTGAAGAGACAGAGGAGGATGCG----AGAGTTTACCTGGGGA 3339
QY 1261 ACAAAACAGGATCTTTTTCGCCCCCTGCTCCAGTNCAGTGTGCTGNAACCCGCTTGG 1320
Db 3340 ACAAAACAGGATCTTTTTCGCCCCCTGCTCCAGT--CGAGTTGGCTTGA--CCGCTTG 3395
QY 1321 ANTACGTACATTTTGTGGGAGNAGCAGGGGAGCAGCTTCCAGCCCTGGGTGAGAAAGG 1380
Db 3396 GATCAGTGACCATTTTGTGGGAGA--CAGGGGAGCAGCTTCCAGCCCTGGGTGAGAAAGG 3454
QY 1381 GTGGCGAGCCCTTCGGGCCCTCACCTNCCAGGCTGCTGTGAGAGTGTCAAGTGTGTA 1440
Db 3455 GTGGCGAGCCCTTCGGGCCCTCACCT--CAGGCTGTCTGTG--AGAGTGTCAAGTGTGTA 3512
QY 1441 AGGNCACCAANCTCAGGNTTCAGTGCAGAACCCAGGTNACAGAGTATGCCCGCCCGNTA 1500
Db 3513 AGGGCCCAAACTC---AGGTTCAAGTGCAGAACCCAGGT--CAGCAGGTATGCCCGCCGTAG 3568
QY 1501 GGTTAANNNGGGGCTCTNAAACCCCTTGCCTNGCCCTNACCTNGCCAGCTCANCC 1560
Db 3569 GTTAA----GGGGGCTCTAAACCCCTTGCCTNGCCCTCA---CCTGGCCAGCTCA--CC 3620
QY 1561 CTTTGGGTGAGGGGAAGAATGCTGACCCCTGGGAAGGCTCCCTGGTGAAGATACAC 1620
Db 3621 CTTTGGGTGAGGGGAAGAATGCTGACCCCTGGGAAGGCT--CCCTGGTGAAGATACAC 3679
QY 1621 CACACTTTTCAAGTTTGTGCAACACAGGCTCTGAGTTGACCTCTCTGGTTACGCAAGGACC 1680
Db 3680 CACACTTTTCAAGTTTGTGCAACACAGGCTCTGAGTTGACCTCTCTGGTTACGCAAGGACC 3739
QY 1681 AAAGAGGTGCTAAGTGAAGTGGTTCTCAGTNCACGATGCCAGATGCCCCCTTGTGCTG 1740
Db 3740 AAAGAGGTGCTAAGTGAAGTGGTTCTCAGT--CCCCAGCATGTGCCCTTTGTGCTG 3798
QY 1741 GCTTACCACTTTCCCCCAGAGCAGAGCCCGGAGCCCTTTCAGGCCCCAGCACTGCCCCAG 1800
Db 3799 GCTTACCACTTTCCCCCAGAGCAGAGCCCGGAGCCCTTTCAGGCCCCAGCACTGCCCCAG 3858
QY 1801 ACTCGCTGGCACTCAGTCCCTCATCTGTAAAGGTGAAGGGTGTGAGGATATGCTCTGA 1860

Db 3859 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGATATGAGGATATGCTCTGA 3918
QY 1861 CAGGAAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGACG 1920
Db 3919 CAGGAAACAGTCTGTGGATGGACATGATCAGTGTCTNAA--GGAAGCAGCAGAGAGAGCG-- 3975
QY 1921 TCCGGCGCCCCAGNCCCACTNATCAGTGTNCCAGCGTGTCTNGGTTNCCCGAGNAGAC 1980
Db 3976 TCCGGCGCCCCAGNCCCACTNATCAGTGT---CCAGCGTGTGTTCCCGAGAGCACAGC-- 4031
QY 1981 GCTNCAAGCATCANCACTGACATNCACTTNGCCCTGCCCCCTNGGCCANGAGAGTCTG 2040
Db 4032 -----TCAGCATCACACTGACATCACCCTGCCCCCTGCCCCCTG--CCAGAGGTTACTG 4082
QY 2041 CCGNACCGCACTTTGCACTNCTGATGNAACCTTCAAGCACCTTTCATGTCTNGCCCTCTNNG 2100
Db 4083 CCG--ACGGCACTTTGCACT--CTGATGACCTTCAAGCACCTTTCATGTCTNGCCCTCTGCA 4139
QY 2101 GCAGGNCAGGNCAGGNCAGTGCANCTGTAGNAGCATANGCAANGCCAGAGATGG 2160
Db 4140 GGGCAGGGCAGGGCAGTGACACT-----GTAGGAGCATAGCAAGCCAGAGATGG 4189
QY 2161 GGTNAAAGGAGNACAGTCTTGTAGCTGTCCANCATGATGTGATNCTTCAAACTCTTCTN 2220
Db 4190 GGTGAA--GGGACACAGTCTTGTAGCTGTCCA--CATGATGTGACT--CCTCAAACTCTT- 4244
QY 2221 NCCAGNATTTCTTAAAGATAGCANCCCTTNGCCCTTNGCCCTTNGCCCTTNGCCCTTCT 2280
Db 4245 --CCAGATTTCTTAAAGATAGCACCCCT--TTCCCATGTCCTCAGCTTNGCCCTTCT 4300
QY 2281 CCCAGGGAGTANCTCAGGATCAGCTCAGCTAGCATTAATCAGCTGTGNAATCGTCAGGGG 2340
Db 4301 CCCAGGGAGCTA--CTCAGGACTCAGCTAGCATTAATCAGCTGTG--AATCGTCAGGGG 4358
QY 2341 TGTCTGTAGCTCAACCTCTGCGGAGGAGGAGCGCGAGACTCGGTGGGAGAGCTCAT 2400
Db 4359 TGTCTGTAGCTCAACCTCTGCGGAGGAGGAGCGCGAGACTCGGTGGGAGAGCTCAT 4418
QY 2401 TCCACATCTTGCACAGACAGCTTTNGTCCAGCTGTCCACATTTGAGTCAGACTGCTCCC 2460
Db 4419 TCCACATCTTGCACAGACAGCTTT--GTCCAGCTGTCCATTTGAGTCAGACTGCTCCC 4477
QY 2461 GGGGAGAGAGCCCCCGCCCCCAGACATATAAGAACTGCGCTTGGTACTGAGAGTCTG 2520
Db 4478 GGGGAGAGAGCCCCCGCCCCCAGACATATAAGAACTGCGCTTGGTACTGAGAGTCTG 4537
QY 2521 GGTGTAGAGACTCTTTGTAAAGCAATAAAGTTTGGGGTGTGACAAATGTTAAAAAA 2579
Db 4538 GGTGTAGAGACTCTTTGTAAAGCAATAAAGTTTGGGGTGTGACAAATGTTAAAAAA 4596

RESULT 5
PCT-US01-08631-15294
; Sequence 15294, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540, 217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649, 167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 15294
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR

LOCATION: (1117)..(1581)
OTHER INFORMATION: 100% homologous to Homo sapiens sperm protein 10, accession
; OTHER INFORMATION: number M82968, Smith-Waterman Score=809.
PCT-US01-08631-15294

Query Match 70.3%; Score 1850.2; DB 1; Length 4596;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

Qy	1	CCCCCTCTCACAGCCAGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGT	60
Db	2100	CCCCCTCTCACAGCCAGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGT	2159
Qy	61	CTGAGGCGGAGCTGGGAGGAGAGTGAACCGGCACTACAGCAAGTGGAGGTCTGAAGA	120
Db	2160	CTGAGGCGGAGCTGGGAGGAGAGTGAACCGGCACTACAGCAAGTGGAGGTCTGAAGA	2219
Qy	121	GCCCTTGGAGGGGAGAAATATAAGCAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC	180
Db	2220	GCCCTTGGAGGGGAGAAATATAAGCAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC	2279
Qy	181	ACAGACCTCTCATGCCAGCCGAGAGAGCTTTCCGCAAGGGGCCCCAGGGCCCGGCCAG	240
Db	2280	ACAGACCTCTCATGCCAGCCGAGAGAGCTTTCCGCAAGGGGCCCCAGGGCCCGGCCAG	2339
Qy	241	CTGAGGAGACAAGGAGAGAGCCCTTAAGTCTCAGCCTCTCTCCACAGAGCCCCCAG	300
Db	2340	CTGAGGAGACAAGGAGAGAGCCCTTAAGTCTCAGCCTCTCTCCACAGAGCCCCCAG	2399
Qy	301	AGCCAAACAGTCTCTCCCTTGAATTTGACGAGGAGGAGTCTGGAGTGTGGAAACCT	360
Db	2400	AGCCAAACAGTCTCTCCCTTGAATTTGACGAGGAGGAGTCTGGAGTGTGGAAACCT	2459
Qy	361	TACCTCTGTCTCTCGAGGAGAGCCCTCTGCAGAAACCCAGCTCACAGAGCGGAAAG	420
Db	2460	TACCTCTGTCTCTCGAGGAGAGCCCTCTGCAGAAACCCAGCTCACAGAGCGGAAAG	2519
Qy	421	CAACCGTCCGAGAGAGAGTGCAGAGCTGGAAATAGAAATATTCTCTCAACAGCCTGT	480
Db	2520	CAACCGTCCGAGAGAGAGTGCAGAGCTGGAAATAGAAATATTCTCTCAACAGCCTGT	2579
Qy	481	CCAGACATTTCTCTGGAGGAGGAGGAGCAATTTCTCTGTCCTCAGCATGACAGCC	540
Db	2580	CCAGACATTTCTCTGGAGGAGGAGGAGCAATTTCTCTGTCCTCAGCATGACAGCC	2639
Qy	541	TCTCCCTGTGGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGGCTCGCGGACA	600
Db	2640	TCTCCCTGTGGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGGCTCGCGGACA	2699
Qy	601	CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTCCAGCTGGA	660
Db	2700	CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTCCAGCTGGA	2759
Qy	661	ACATGCTGTGGCCGGGGGGGGCCCAACGACACCCCAAGCTATTCAATGGTGAAG	720
Db	2760	ACATGCTGTGGCCGGGGGGGGCCCAACGACACCCCAAGCTATTCAATGGTGAAG	2819
Qy	721	TCCAAATACAGTCTTTAATGGTGAACACCTGCACATCCCGGAGTTCACCGGGTCAAG	780
Db	2820	TCCAAATACAGTCTTTAATGGTGAACACCTGCACATCCCGGAGTTCACCGGGTCAAG	2879
Qy	781	TGGGAGACATCGGCATCGAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCA	840
Db	2880	TGGGAGACATCGGCATCGAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCA	2939
Qy	841	CCAAAGACGGGAGCTGTTCGCTTCAACATGAGAGTGCAGACTCGGGCATCGACTGC	900
Db	2940	CCAAAGACGGGAGCTGTTCGCTTCAACATGAGAGTGCAGACTCGGGCATCGACTGC	2999
Qy	901	AGTGCACTGGCCCTGTGAGGAGCTTCGCTGAGCTGGAGGTCAGCATGGCCAG	960
Db	3000	AGTGCACTGGCCCTGTGAGGAGCTTCGCTGAGCTGGAGGTCAGCATGGCCAG	3059

Qy	961	TGAGAAACAGGCGCTTAACCTGCGCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT	1020
Db	3060	TGAGAAACAGGCGCTTAACCTGCGCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT	3119
Qy	1021	TCCTGCTCGGTGACGATGCTGCCCTGAAAACACAGCTCAGCGTTCCAGGGGATYTG	1080
Db	3120	TCCTGCTCGGTGACGATGCTGCCCTGAAAACACAGCTCAGCGTTCCAGGGGATYTG	3177
Qy	1081	NCAGAGCCCCCGGCTCARCAGNTGGGAAACAGGCGCTTCGNACAGNAGGTTNGGGG	1140
Db	3178	GCAGAGCCCCCGGCTCA--CAGTGGGAAACAGGCGCTCG----CAGCAGCAAGTGGG	3230
Qy	1141	CAAGCAGAGATGCTCCAGGATTTTACANCTTACAGCCNTGAGCCNTGCCANCCCTGCTGA	1200
Db	3231	GCAAGCAGAGATGCTCCAGGATTTTACACCTGAGCCCTGCCAC-----CCTGCTGA	3284
Qy	1201	AAACATVNCGCCACGCTGAAGACAGAGAGGATGAGGATGAGGATGAGGATGAGGATG	1260
Db	3285	AAACATVNCGCCACGCTGAAGACAGAGAGGATGAGGATGAGGATGAGGATGAGGATG	3340
Qy	1261	ACAAACAGGAGTCTTTTCTGCCCCCTGCTCCAGTNCAGATTTGGCCCTGNACCCGCTTG	1320
Db	3341	ACAAACAGGAGTCTTTTCTGCCCCCTGCTCCAGT--CGAGTTGGCTGA--CCCGCTTG	3396
Qy	1321	ANTCAGTGACATTTGTTGGCAGANAGGGGAGAGAGCTTCCAGCTGGGTGAGAGGG	1380
Db	3397	GATCAGTGACATTTGTTGGCAG--CAGGGGAGAGAGCTTCCAGCTGGGTGAGAGGG	3455
Qy	1381	GTGGGCGAGCCCTCGGCCCCCTCACCTNCCAGGCTGCTGTGNAGAGTGTCAAGTGTGA	1440
Db	3456	GTGGGCGAGCCCTCGGCCCCCTCACCT--CCAGGCTGCTGTG--AGAGTGTCAAGTGTGA	3513
Qy	1441	AGGNCGCCAAANTCTCAGGNTTTCAGTGAGAAACAGGTCNCCAGAGGATGCGCCGCCNTA	1500
Db	3514	AGGNCGCCAAATC--AGGTTTCAGTGAGAAACAGGTCNCCAGAGGATGCGCCGCCNTA	3569
Qy	1501	GGTTAANGGGGCGCTCTNAAACCCCTTGCCCTNGGCTNCACTNGGCGAGCTCANCCC	1560
Db	3570	GGTTA--GGGGCGCTCTNAAACCCCTTGCCCTGCGCTCA--CCTGCGCAGCTCA--CCC	3621
Qy	1561	CTTTTGGGTGAGGGGAAAGATGCTGACCTGGGAGGCTTCCCTGCTGAGATATACAC	1620
Db	3622	CTTTTGGGTGAGGGGAAAGATGCTGACCTTGGGAGGCT--CCTGCTGAGATATACAC	3680
Qy	1621	CACATTTTCAGGTGTTGCAACACAGGTCCTCAGTTGACCTCTGTTTCAGCAGGAGCC	1680
Db	3681	CACATTTTCAGGTGTTGCAACACAGGTCCTCAGTTGACCTCTGTTTCAGCAGGAGCC	3740
Qy	1681	AAAGAGGTGTGTAAGTGAAGTGTCTCAGTNCGCCAGACATGTGCCCTTTGCTGCTG	1740
Db	3741	AAAGAGGTGTGTAAGTGAAGTGTCTCAGT--CCCAGACATGTGCCCTTTGCTGCTG	3799
Qy	1741	GCTACACATCTTCCCAGAGCAGAGCCCGCCGAGCCCTTTCAGGCGCAGACTGCCCCAG	1800
Db	3800	GCTACACATCTTCCCAGAGCAGAGCCCGCCGAGCCCTTTCAGGCGCAGACTGCCCCAG	3859
Qy	1801	ACTCGCTGGCAGTTCCTCATCTGTAAAGGTGAAGGTGATGAGGATATGCTGTA	1860
Db	3860	ACTCGCTGGCAGTTCCTCATCTGTAAAGGTGAAGGTGATGAGGATATGCTGTA	3919
Qy	1861	CAGGAAACAGTCTGTGGATGGACATGATGCTGCTNAGGAAAGCAGCAGAGAGAGAG	1920
Db	3920	CAGGAAACAGTCTGTGGATGGACATGATGCTGCTNAGGAAAGCAGCAGAGAGAGAG	3976
Qy	1921	TCCGGCGCCCGCAGNCCCACTNATAGTGTNCCAGGCTGCTGCTNCCCGCAGNAGACA	1980
Db	3977	TCCGGCGCCCGCAGNCCCACTNATAGTGTNCCAGGCTGCTGCTNCCCGCAGNAGACA	4032
Qy	1981	GCTNAGNATCANCACTGACACTNCCAGCTTCCCTGCGCTGCGCTGCGCTGCGCTG	2040
Db	4033	-----TACGATCACTGACACTCACTCCCTGCGCTGCGCTGCGCTGCGCTGCGCTG	4083
Qy	2041	CCGACGCGACTTTGACACTTCTGATGNACTTCAAGACCTTTTCTGCTGCTGCTGCTG	2100

```

Db 4084 CCG-ACGGCACTTGGCACT--CTGATGACCTCAAGCACATTTTCATGGTGCCTCTCGCA 4140
Qy 2101 GGAGGNCAGGNCAGGNCAGTGCACNCTGTAGNAGCATANGAAGCCAGAGATGG 2160
Db 4141 GGGCAGGCGCAGGCGAGTGCACACT-----GTAGGAGCATAGCAAGCCAGAGATGG 4190
Qy 2161 GGTGNAAGGNCACAGCTTTGAGCTGTCANCAATGATGTGACTTNCCTCAAACTCTTN 2220
Db 4191 GGTGAA--GGACACAGCTTTGAGCTGTCCA-CATGATGTGACT-CCTCAAACTCTT- 4245
Qy 2221 NCCAGNATTTCTTAAGATAGCANCCTTTNCCCAATTCGCCAGCTTAGCTCTTCT 2280
Db 4246 --CCAGATTTCTTAAGATAGCACCCCTT--TTCCCAATTCGCCAGCTTAGCTCTTCT 4301
Qy 2281 CCCAGGAGCTTANTCAGGACTCAGCTAGCATTAATCAGCTGTGNAATCGTCAGGGG 2340
Db 4302 CCCAGGAGCTA-CTCAGGACTCAGCTAGCATTAATCAGCTGTG- AATCGTCAGGGG 4359
Qy 2341 TGTCTGCTAGCTCAACCTCTTGGGCGAGGAGCGCCGAGACTCCGTGGGAGAGCTCAT 2400
Db 4360 TGTCTGCTAGCTCAACCTCTTGGGCGAGGAGCGCCGAGACTCCGTGGGAGAGCTCAT 4419
Qy 2401 TCCCACTCTTCCCAAGACAGCTTTNGTCCAGCTGTCCACATTAAGTCAAGCTGCTCC 2460
Db 4420 TCCCACTCTTCCCAAGACAGCTTT- GTCCAGCTGTCCACATTAAGTCAAGCTGCTCC 4478
Qy 2461 GGGGAGAGAGCCCGGCGCCCGAGCATTAAGAACTGAGCTTGGTACTGCAGAGCTG 2520
Db 4479 GGGGAGAGAGCCCGGCGCCCGAGCATTAAGAACTGAGCTTGGTACTGCAGAGCTG 4538
Qy 2521 GCTGTGAGAGACTCTTTGTAGCAATAAAGTTTGGGGTGTAGCAAAATGTTAAAAA 2578
Db 4539 GGTGTGAGAGAACTCTTTGTAGCAATAAAGTTTGGGGTGTAGCAAAATGTTAAAAA 4596

RESULT 6
PCT-US02-18947-975
; Sequence 975, Application PC/TUS0218947
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-228
; CURRENT APPLICATION NUMBER: PCT/US02/18947
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 975
; TYPE: DNA
; LENGTH: 4596
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_003954
; DATABASE ENTRY DATE: 2001-06-18
PCT-US02-18947-975

Query Match 70.3%; Score 1850.2; DB 1; Length 4596;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

Qy 1 CCCTCTCAGCGCCAGGCGCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGT 60
Db 2100 CCCTCTCAGCGCCAGGCGCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGT 2159
Qy 61 CTGACGGAGCTGGAGGAGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120
Db 2160 CTGACGGAGCTGGAGGAGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 2219
Qy 121 GCCTTTGGGGGAGAGATATAAGAACCAAGACATCCCGCCAAATCAAGCAATACC 180
Db 2220 GCCTTTGGGGGAGAGATATAAGAACCAAGACATCCCGCCAAATCAAGCAATACC 2279
```

```

Qy 181 ACCAGACCCCTCCATGCTCCAGCCAGAGAGCTTTTCGCAAGGGCCCCAGGGCCCCGGCCAG 240
Db 2280 ACCAGACCCCTCCATGCTCCAGCCAGAGAGCTTTTCGCAAGGGCCCCAGGGCCCCGGCCAG 2339
Qy 241 CTGAGGAGACAAAGCAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACACAGAGCCGCCAG 300
Db 2340 CTGAGGAGACAAAGCAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACACAGAGCCGCCAG 2399
Qy 301 AGCCAAACAAAGTCTCTCTCCCTTCTGACTTTGAGCAAGGAGGAGTCTGGATGTGGAAACCT 360
Db 2400 AGCCAAACAAAGTCTCTCTCCCTTCTGACTTTGAGCAAGGAGGAGTCTGGATGTGGAAACCT 2459
Qy 361 TACTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Db 2460 TACTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2519
Qy 421 CAACCGTCCCGGAGCAGGAGACTCGACAGCTGGAATATAGAAATATTCTCTCAACAGCTGT 480
Db 2520 CAACCGTCCCGGAGCAGGAGACTCGACAGCTGGAATATAGAAATATTCTCTCAACAGCTGT 2579
Qy 481 CCCAGCCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 2580 CCCAGCCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2639
Qy 541 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Db 2640 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2699
Qy 601 CCCTGAGCTCAGGCGTACACTCTCTGGAGAGCAGGAGCCGAGGCTCGAAGCTTCCAGCTTGA 660
Db 2700 CCCTGAGCTCAGGCGTACACTCTCTGGAGAGCAGGAGCCGAGGCTCGAAGCTTCCAGCTTGA 2759
Qy 661 ACATGTGTCTGGCCCGGGGGCGGCCCAACGACATCCCAAGCTATTTCAATGGTGTGAAG 720
Db 2760 ACATGTGTCTGGCCCGGGGGCGGCCCAACGACATCCCAAGCTATTTCAATGGTGTGAAG 2819
Qy 721 TCCAAATACAGTCTCTTAATGGTGAACATCTGACATCCCGGAGTTCACCGGCTCAAG 780
Db 2820 TCCAAATACAGTCTCTTAATGGTGAACATCTGACATCCCGGAGTTCACCGGCTCAAG 2879
Qy 781 TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTGAGCTTCCAGCTTGGTCA 840
Db 2880 TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTGAGCTTCCAGCTTGGTCA 2939
Qy 841 CCAAAGACGGGCGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 2940 CCAAAGACGGGCGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2999
Qy 901 AGTGACACTGGCCCGCTGATGGCAGCTTCCGCTGGAGCTGGAGGCTCAAGCATGGCCAGC 960
Db 3000 AGTGACACTGGCCCGCTGATGGCAGCTTCCGCTGGAGCTGGAGGCTCAAGCATGGCCAGC 3059
Qy 961 TGGAGAACAGGCGCTTAACCTCGCCCTCCACCGCGGGTCCACACTGCGCGGAAAGCAGCT 1020
Db 3060 TGGAGAACAGGCGCTTAACCTCGCCCTCCACCGCGGGTCCACACTGCGCGGAAAGCAGCT 3119
Qy 1021 TCCTGTCTGGTGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 3120 TCCTGTCTGGTGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3177
Qy 1081 NCCAGCCCCCGGCTCAGCAGTGGGAGCAGGCTCGCAGGCTCGCAGGCTCGCAGGCTCGCAGG 1140
Db 3178 GCCAGCCCCCGGCTCAGCAGTGGGAGCAGGCTCGCAGGCTCGCAGGCTCGCAGGCTCGCAGG 3230
Qy 1141 CAAGCNAAGATGCTCTCCAGGATTTTCAACGCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAG 1200
Db 3231 GCAAGCNAAGATGCTCTCCAGGATTTTCAACGCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAG 3284
Qy 1201 AAACATYTTCCCGCAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 3285 AAACATCATCCGACAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3340
Qy 1261 ACAAAACAGGAGATCTTTTCTCTGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320
```


Qy	541	TCTCCCTGTCGGATGACAGTGAAGAAGACCCATCAAAAGGCCTCTCAAAGCTCGCGGACA	600
Db	2640	TCTCCCTGTCGGATGACAGTGAAGAAGACCCATCAAAAGGCCTCTCAAAGCTCGCGGACA	2699
Qy	601	CCCTGAGCTCAGGGGTACACTCTCGGAGCAGCAGGCGGAGGCTCGAAGCTTCAGCTTGA	660
Db	2700	CCCTGAGCTCAGGGGTACACTCTCGGAGCAGCAGGCGGAGGCTCGAAGCTTCAGCTTGA	2759
Qy	661	ACATGGTGTGGCCCGGGGGCGGCCACCCAGCACACCCCAAGCTATTTCAATGGTGTGAAG	720
Db	2760	ACATGGTGTGGCCCGGGGGCGGCCACCCAGCACACCCCAAGCTATTTCAATGGTGTGAAG	2819
Qy	721	TCCAAATACAGTCTCTTTAATGGTGAAACACTGTGCAATCCGCGGAGTTCACACCGGCTCAAAG	780
Db	2820	TCCAAATACAGTCTCTTTAATGGTGAAACACTGTGCAATCCGCGGAGTTCACACCGGCTCAAAG	2879
Qy	781	TGGGAGACATCGCCACTTGGCATCAGCAGCAGAGATCCACAGCTGCAGCTTCAGCTTGGTCA	840
Db	2880	TGGGAGACATCGCCACTTGGCATCAGCAGCAGAGATCCACAGCTGCAGCTTCAGCTTGGTCA	2939
Qy	841	CCAAAGACGGCAGCCTGTTCCGTACGACATGGAGGTGCCAGACTCGGGGATCGACCTGC	900
Db	2940	CCAAAGACGGCAGCCTGTTCCGTACGACATGGAGGTGCCAGACTCGGGGATCGACCTGC	2999
Qy	901	AGTGCACACTGGCCCTTGATGGCAGCTTCGCCTGGAGCTGGAGGGTCAAGCATGGCCAGC	960
Db	3000	AGTGCACACTGGCCCTTGATGGCAGCTTCGCCTGGAGCTGGAGGGTCAAGCATGGCCAGC	3059
Qy	961	TGGAGAACAGGCCCTTAACCTTGCCCTTCCACGGCCGGCTCCACAGCTGCGGAAGACAGCCT	1020
Db	3060	TGGAGAACAGGCCCTTAACCTTGCCCTTCCACGGCCGGCTCCACAGCTGCGGAAGACAGCCT	3119
Qy	1021	TCCTGCTCGGTGACAGATGCTGCCCTCGAAACACAGGCTCAGCCGTTCCCGAGGGGATYTG	1080
Db	3120	TCCTGCTCGGTGACAGATGCTGCCCTCGAAACACAGGCTCAGCCGTTCCCGAGGGGATYTG	3177
Qy	1081	NCCAGCCCCCGGCTCARGAGTGGGAACCAGGGCCCTCGNCAGCAGCNAAAGTNGGGGG	1140
Db	3178	GCCAGCCCCCGGCTCA - CAGTGGGAACCAGGGCCCTCG - - - - - CAGCAGCAAGGTGGGG	3230
Qy	1141	CAAGCNAGATGCCTCCAGGATTTACANCTTGAGCCCTGAGCCCTGCCCCANCCCTGCTGAADA	1200
Db	3231	GCAAGCAGAAATGCCCTCCAGGATTTACACTGAGCCCTCGCCAC - - - - - CCTGCTGA	3284
Qy	1201	AAACAYTNCGCCACGCTGAAGAGACAGAGAGATGNCAGGAGTTNNACCTYGGGAA	1260
Db	3285	AAAAACATCGCCACGTGAAGAGACAGAGAGATGNCAGGAGTTNNACCTYGGGAA	3340
Qy	1261	ACAAACAGGGATCTTTNTTCTGCCCTTGCTCCAGTNCGAGTGTGGCCTGNACCCGCTTGG	1320
Db	3341	ACAAAACAGGGATCTTTT - TTCTGCCCTTGCTCCAGT - CGAGTTGGCCTGA - - - - - CCGGCTTG	3396
Qy	1321	ANTCAGTGACCATTTGTTGGCAGANCAGGGGAGAGCAGTTCACAGCTCGGGTCAGAAGGG	1380
Db	3397	GATCAGTGACCATTTGTTGTCAGAG - CAGGGGAGAGCAGTTCACAGCTCGGGTCAGAAGGG	3455
Qy	1381	GTGGGCGAGCCCTTCGCGCCCTCACCTNCCAGCTGCTGTGNAGAGTGTCAAGTGTGA	1440
Db	3456	GTGGGCGAGCCCTTCGCGCCCTCACCT - - - - - CAGGCTGCTGTG - AGAGTGTCAAGTGTGA	3513
Qy	1441	AGGNCNCAANCTCAGNTTTCAGTGAGAACACAGGTNCCAGGATGTCGCCGCCGNTA	1500
Db	3514	AGGSCCCAAATC - - - - - AGGTTTCAGTGAGAACACAGGT - CAGCAGGTATGCCGCCGCTAG	3569
Qy	1501	GGTTAANNGGGGCCCTCTNAAACCCCTTGCTNNGGCTNCACTNNGGCCAGCTCANCCC	1560
Db	3570	GTTTAA - - - - - GGGGGCCCTCTAAAACCCCTTGCTGGCCCTCA - - - - - CCTGGCCAGCTCA - CCC	3621
Qy	1561	CTTTTGGGTGTAGGGGAAGAAATGCTGACCTCGAGGAGGCTWCCTTGTTGAGTAACAC	1620
Db	3622	CTTTTGGGTGTAGGGGAAGAAATGCTGACCTCGAGGAGGCT - CCTGGTAGAATAACAC	3680
Qy	1621	CACACTTTTTCAGGTGTGTTGCAACACAGGCTCTCAGTGTGACCTCTGTTTCAGCCAAAGGACC	1680

Db	3681	CACACTTTTCAGGTGTGTGCAACACAGAGTCTCTGAGTTGACCTCTGTTT CAGCCAGGACC 3740
Qy	1681	AAAGAAGGTGTGAAGTGAAGTGGTTCTCAGTNCCCACAGACATGTGCCCCCTTTGCTGCTG 1740
Db	3741	AAAGAAGGTGTGAAGTGAAGTGGTTCTCAGT--CCCCACATGTGCCCCCTTTGCTGCTG 3799
Qy	1741	GCTACCACTCTTCCACAGACACAGCCCGAGCCCTTCAGCCCCAGCACTGCCCCAG 1800
Db	3800	GCTACCACCTTCTCCACAGACACAGCCCGAGCCCTTCAGCCCCAGCACTGCCCCAG 3859
Qy	1801	ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGTGAAGGGTGATCAGAGATATGCTGA 1860
Db	3860	ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGTGAAGGGTGATCAGAGATATGCTGA 3919
Qy	1861	CAGAAACAGTCTCTGGATGACATGATCAGTGTCTNAAGGNAAGACAGACAGAGACGY 1920
Db	3920	CAGAAACAGTCTCTGGATGACATGATCAGTGTCTAA--GGAAAGCAGACAGAGAGACG- 3976
Qy	1921	TCGGGCCCCCAGNCCCCACTNATCAGTGTGTCAGCGTGTCTGTTGCCAGNACACA 1980
Db	3977	TCGGGCCCCCAGCCCCCACTATCAGTGT---CCAGCGTGTGTTCTCCACAGACACAGC- 4032
Qy	1981	GCTNCAGNATCANCACTGACACTNCACCTGSCCTGCCCTTGGCCANGAGGGTACTG 2040
Db	4033	-----TCAGCATCACACTGACACTCACCTGCCCTGCCCTGG--CCAGAGGGTACTG 4083
Qy	2041	CCGNA CGGCACCTTTGCA CMTCTGATGNACCTCAAAGCACTTTTCATGGCTNGCCCTCTNG 2100
Db	4084	CCG-ACGGCACTTTTGCACT--CTGATGACCTCAAAGCACTTTTCATGGCTGCCCTCTGGCA 4140
Qy	2101	GCAGGNCAGGNCAGGNCAGGNCAGTGACANCTGTAGGNAGCATANGCAAGCCAGGATGG 2160
Db	4141	GGGCAGGGCAGGCGAGTGACACT-----GTAGGAGCATAGCAAGCCAGGATGG 4190
Qy	2161	GGTGNAAGGANCA CAGTCTTTGAGCTGTCCANCATGCACTGTGACTNCCTCAAACCTCTTN 2220
Db	4191	GGTGAA--GGGACACAGTCTTGAGCTGTCCA-CATGCATGTGACT--CCTCAAACCTCTT- 4245
Qy	2221	NCCAGNATTTCTTAAGAATPAGCANCCCCCTTNNCCCACTTGCCCCCAGCTTAGCCTCTTCT 2280
Db	4246	--CCAGATTTCTCTAAGAATPAGCACCCCC--TTCCCACTTGCCCCCAGCTTAGCCTCTTCT 4301
Qy	2281	CCCAGGGAGCTANCTCAGGACCTCAGTGACATTAATCAGCTGTGNAATCGTCAAGGGG 2340
Db	4302	CCCAGGGAGGCTA-CTCAGGACCTCAGTGACATTAATCAGCTGTG-AACTGTCAGGGGG 4359
Qy	2341	TGTCTGTAGCCTCAACCTCCTGGGGCAGGGGACGCCGAGACTCCGCTGGGAGAGCTCAT 2400
Db	4360	TGTCGTAGCCTCAACCTCTGGGCAGGGGACGCCGAGACTCCGCTGGGAGAGCTCAT 4419
Qy	2401	TCCCACTATTGCGACAGACGCTTTNGTCCAGCTGTCCAATTGATGACAGCTGCTCCC 2460
Db	4420	TCCCACTATTGCGCAAGACAGCCTTTT-GTCCAGCTGTCCAATTGATGACAGCTGCTCCC 4478
Qy	2461	GGGAGAGAGCCCCGGCCCCCAGCATAAAGAACTGCAGCCCTTGTGCTGCAGAGTCTG 2520
Db	4479	GGGAGAGAGCCCCGGCCCCCAGCATAAAGAACTGCAGCCCTTGTGCTGCAGAGTCTG 4538
Qy	2521	GGTTGTAGAACTCTTTGTAAAGCAATAAGTTTGGGGTGATCAAAATGTTAAAAA 2578
Db	4539	GGTTGTAGAACTCTTTGTAAAGCAATAAGTTTGGGGTGATCAAAATGTTAAAAA 4596

```

RESULT 8
US-09-053-375B-1234
; Sequence 1234, Application US/09053375B
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Bibilashvili, Robert
; TITLE OF INVENTION: Nucleic Acid Arrays
; FILE REFERENCE: CLON-006
; CURRENT APPLICATION NUMBER: US/09/053,375B

```

; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 1543
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1234
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-053-375B-1234

Query Match 70.3%; Score 1850.2; DB 16; Length 4596;

Best Local Similarity 90.5%; Pred. No. 0;
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

Qy	1	CCCTCTCACAGCCACGCGCCATCCAAAGAGGGCTTGAGGAAAGAGCCCATCACCGCTGT	60
Db	2100	CCCTCTCACAGCCACGCGCCATCCAAAGAGGGCTTGAGGAAAGAGCCCATCACCGCTGT	2159
Qy	61	CTGACGGGAGCTGGGAGGAGGTGAACCGGCACTACAGCAAGTGGGAGTCTGAAGA	120
Db	2160	CTGACGGGAGCTGGGAGGAGGTGAACCGGCACTACAGCAAGTGGGAGTCTGAAGA	2219
Qy	121	GCCCTTGAGGGGAGATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC	180
Db	2220	GCCCTTGAGGGGAGATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC	2279
Qy	181	ACCAGACCTTCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCCAGGGCCCCGGCCAG	240
Db	2280	ACCAGACCTTCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCCAGGGCCCCGGCCAG	2339
Qy	241	CTGAGGAGACACAGGACAGGCCCCCTAAGCTCCAGGCTCTCTCCACAGAGCCCCCAG	300
Db	2340	CTGAGGAGACACAGGACAGGCCCCCTAAGCTCCAGGCTCTCTCCACAGAGCCCCCAG	2399
Qy	301	AGCCAAACAACTCTCCCTTGACTTTTGACAGAGGAGGAGTCTGGGATGTGGGAACCT	360
Db	2400	AGCCAAACAACTCTCCCTTGACTTTTGACAGAGGAGGAGTCTGGGATGTGGGAACCT	2459
Qy	361	TACCTCTGTCTCTCTGGAGCCAGCCCTCTGCGAAGAACCCAGCTCCACAGAGCGGAAAG	420
Db	2460	TACCTCTGTCTCTCTGGAGCCAGCCCTCTGCGAAGAACCCAGCTCCACAGAGCGGAAAG	2519
Qy	421	CAACGCTCCGGAGCAGGAACTGACAGAGCTGGGAAATAGAAATTCTCTCAACAGCCTGT	480
Db	2520	CAACGCTCCGGAGCAGGAACTGACAGAGCTGGGAAATAGAAATTCTCTCAACAGCCTGT	2579
Qy	481	CCAGACCAATTTCTCTGAGAGCAGGAGCAAAATCTCTGCTGCTCAGCATCGACAGCC	540
Db	2580	CCAGACCAATTTCTCTGAGAGCAGGAGCAAAATCTCTGCTGCTCAGCATCGACAGCC	2639
Qy	541	TCTCCCTGTCTGGATGACAGTGAAGAAACCCATCAAAAGGCTCTCAAAAGCTTCGGGGACA	600
Db	2640	TCTCCCTGTCTGGATGACAGTGAAGAAACCCATCAAAAGGCTCTCAAAAGCTTCGGGGACA	2699
Qy	601	CCCTGAGCTCAGGCGTACACTCTGGAGACGACAGCCGAGGCTCGAAGCTCCAGCTGGA	660
Db	2700	CCCTGAGCTCAGGCGTACACTCTGGAGACGACAGCCGAGGCTCGAAGCTCCAGCTGGA	2759
Qy	661	ACATGGTGTGGCCCGGGGGGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG	720
Db	2760	ACATGGTGTGGCCCGGGGGGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG	2819
Qy	721	TCCAAATACAGTCTCTTAATGGTGAACAACCTTGACATCCGGGAGTTCACCGGCTCAAG	780
Db	2820	TCCAAATACAGTCTCTTAATGGTGAACAACCTTGACATCCGGGAGTTCACCGGCTCAAG	2879
Qy	781	TGGAGACATCGCCCATCTGGCATCAGCAGCAGATCCAGCTGACGCTTCAGCTTGTGCA	840
Db	2880	TGGAGACATCGCCCATCTGGCATCAGCAGCAGATCCAGCTGACGCTTCAGCTTGTGCA	2939
Qy	841	CCAAAGACGGCAGCTTCTGCTTACGACATGGAGGTGCCAGCTCGGGCATCGACCTGC	900
Db	2940	CCAAAGACGGCAGCTTCTGCTTACGACATGGAGGTGCCAGCTCGGGCATCGACCTGC	2999

Qy	901	AGTGCACTGGCCCTGTGATGGGAGCTTCCCTGGAGCTGGAGGTCAGCATGGCCAGC	960
Db	3000	AGTGCACTGGCCCTGTGATGGGAGCTTCCCTGGAGCTGGAGGTCAGCATGGCCAGC	3059
Qy	961	TGGAGAACAGCCCTTAACCTTCCCTCCACCGCGGCTCCACACTGCCGGAAGCAGCCT	1020
Db	3060	TGGAGAACAGCCCTTAACCTTCCCTCCACCGCGGCTCCACACTGCCGGAAGCAGCCT	3119
Qy	1021	TCCTGCTCGGTGACGATGCTCCCTGAAAACACAGGCTCAGCGGTTCCAGGGGATGTG	1080
Db	3120	TCCTGCTCGGTGACGATGCTCCCTGAAAACACAGGCTCAGCGGTTCCAGGGGATGTG	3177
Qy	1081	NCAGCCCCCGGCTCARGAGTGGGAAACAGGGCTCTGNCAGNACGNAAAGTNGGGGG	1140
Db	3178	GCCAGCCCCCGGCTCA--CAGTGGGAAACAGGGCTCTG-----CAGCAGCAAGGTGGG	3230
Qy	1141	CAAGCAGAGTGCCTCCAGGATTTACANCTTGAGCCCTGAGCCCTGCCANCCCTGCTGA	1200
Db	3231	GCAAGCAGAGTGCCTCCAGGATTTACACTGAGCCCTGCCCCAC-----CCTGTGA	3284
Qy	1201	AAACATYNNCCGCCACGTGAAGAGACAGAGGAGGATGNCAGAGGATTTNNACCTYGGGAA	1260
Db	3285	AAACATYNNCCGCCACGTGAAGAGACAGAGGAGGATGNCAGAGGATTTNNACCTYGGGAA	3340
Qy	1261	ACAAAACAGGATCTTTNTTCTGCTGCTGTCAGTNCAGTGGCTGNAACCGCTTGG	1320
Db	3341	ACAAAACAGGATCTTTNTTCTGCTGCTGTCAGTNCAGTGGCTGNAACCGCTTGG	3396
Qy	1321	ANTCAGTGACATTTGTTGGCAGNACAGGGAGAGCAGCTTCCAGCTGGTTCAGAGGG	1380
Db	3397	GATCAGTGACATTTGTTGGCAGNACAGGGAGAGCAGCTTCCAGCTGGTTCAGAGGG	3455
Qy	1381	GTGGGCGAGCCCTTCGGCCCTCAACCTTNCAGGCTGCTGTGAGAGTGTCAAGTGTGTA	1440
Db	3456	GTGGGCGAGCCCTTCGGCCCTCAACCTTNCAGGCTGCTGTGAGAGTGTCAAGTGTGTA	3513
Qy	1441	AGGNNCCAAANTCAGGNTTCAGTGCAGAACACAGGTCACAGAGTATGCCCCCGNTA	1500
Db	3514	AGGNNCCAAANTCAGGNTTCAGTGCAGAACACAGGTCACAGAGTATGCCCCCGNTA	3569
Qy	1501	GGTTAANNNGGGGCGCTCTNAAACCCCTTGGCTTGGCTTNCACCTTNGGCGCAGCTC	1560
Db	3570	GGTTAANNNGGGGCGCTCTNAAACCCCTTGGCTTGGCTTNCACCTTNGGCGCAGCTC	3621
Qy	1561	CTTTTGGGTGAGGGGAAAAAGATGCTGACCTCTGGGAGGCTTCCCTGGTAGAATACAC	1620
Db	3622	CTTTTGGGTGAGGGGAAAAAGATGCTGACCTCTGGGAGGCTTCCCTGGTAGAATACAC	3680
Qy	1621	CACACTTTTTCAGGTTGTGCAACACAGGCTCCTGAGTTGACCTCTGTTTCAGCCAGGACC	1680
Db	3681	CACACTTTTTCAGGTTGTGCAACACAGGCTCCTGAGTTGACCTCTGTTTCAGCCAGGACC	3740
Qy	1681	AAAGAGGTGTGTAAGTGAAGTGTCTCAGTNCACAGACATGTCGCCCTTTGCTGCTG	1740
Db	3741	AAAGAGGTGTGTAAGTGAAGTGTCTCAGTNCACAGACATGTCGCCCTTTGCTGCTG	3799
Qy	1741	GCTACCACTCTTCCAGAGCAGCAGCCCGGAGCCCTTTCAGGCCAGACACTGCCCCAG	1800
Db	3800	GCTACCACTCTTCCAGAGCAGCAGCCCGGAGCCCTTTCAGGCCAGACACTGCCCCAG	3859
Qy	1801	ACTCGCTGGCACTCAGTTCCTCTCATCTGTAAGAGTGAAGGTTGATGAGGATATGCTGA	1860
Db	3860	ACTCGCTGGCACTCAGTTCCTCTCATCTGTAAGAGTGAAGGTTGATGAGGATATGCTGA	3919
Qy	1861	CAGGAAACAGTCTGTGATGACATGATCAGTCTNAAAGNAAAGCAGCAGAGAGAGCGY	1920
Db	3920	CAGGAAACAGTCTGTGATGACATGATCAGTCTNAAAGNAAAGCAGCAGAGAGAGCGY	3976
Qy	1921	TCCGGGCGCCCGAGNCCCACTNATCAGTGTNCCAGGCTGCTTNGCTTCCCGAGAGCACA	1980
Db	3977	TCCGGGCGCCCGAGNCCCACTNATCAGTGTNCCAGGCTGCTTNGCTTCCCGAGAGCACA	4032
Qy	1981	GCTNCAAGNCACTCANCACTGACACTNCAACCTTNGCCCTTGGCCCTNGGCGCANGAGG	2040

Db 3285 AAAACATCCCGACCTGAGAGACAGAGAGATGGC-----AGGAGTTACCTGGGGAA 3340
Qy 1261 AAAAAACAGGAGATCTTNTTCTGCCCTGCTCCAGTNCAGATGGCCCTGNAACCGCTGG 1320
Db 3341 AAAAAACAGGAGATCTTT-TTCGCGCCCTGCTCCAGT-CGAGTTGGCCTGA--CCGCTTG 3396
Qy 1321 ANTCAAGTACCATTTGTTGGCAGANAGGGGAGAGAGCTTCAGCCCTGGGTCAAGGG 1380
Db 3397 GATCAGTGAACATTTGTTGGCAGA-CAGGGAGAGAGCTTCAGCCCTGGGTCAAGAGG 3455
Qy 1381 GTGGGAGAGCCCTTCGGGCCCTTCACTCCCTNCCAGGCTGCTGTGNAGAGTGTCAAGTGTGA 1440
Db 3456 GTGGGAGAGCCCTTCGGGCCCTTCACTCCCT-CCAGGCTGCTGTG-AGAGTGTCAAGTGTGA 3513
Qy 1441 AGGNCCTCAANTCAGNTTTCAGTGCAGAACAGGTNCCAGAGTATGCCCGCCGNTA 1500
Db 3514 AGGCCCAAACTC---AGGTTCAAGTGCAGAACAGGT-CAGCAGGTATGCCCGCCGTAG 3569
Qy 1501 GGTAAANNNGGGGCCCTCTNAAACCCCTTGCTNGGCTNCACTNGGCCAGCTCANCCC 1560
Db 3570 GTTAA-----GGGGGCCCTCTAAACCCCTTGCTGGCCCTCA---CCTGGCCAGCTCA-CCC 3621
Qy 1561 CTTTGTGGGTGAGGGGAAAAGATGCTGACCTCTGGGAAGGCTWCCCTGGTGAATACAC 1620
Db 3622 CTTTGTGGGTGAGGGGAAAAGATGCTGACCTCTGGGAAGGCT-CCCTGGTGAATACAC 3680
Qy 1621 CACATTTTTCAGGTTGTTGCAACACAGGTCTGAGTTGACCTCTGTGTTTCAGCCAGGACC 1680
Db 3681 CACATTTTTCAGGTTGTTGCAACACAGGTCTGAGTTGACCTCTGTGTTTCAGCCAGGACC 3740
Qy 1681 AAAGAAGGTGTAAAGTGAAGTGTCTTCAGTNCCTCCAGACATGTGCCCTTTGCTGCTG 1740
Db 3741 AAAGAAGGTGTAAAGTGAAGTGTCTTCAGT-CCCCAGACATGTGCCCTTTGCTGCTG 3799
Qy 1741 GCTACCACTCTTCCCCAGAGCAGAGCCCGAGCCCTTTCAGGCCAGCACTGCCCCAG 1800
Db 3800 GCTACCACTCTTCCCCAGAGCAGAGCCCGAGCCCTTTCAGGCCAGCACTGCCCCAG 3859
Qy 1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGATGACAGTATGCTCTGA 1860
Db 3860 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGATGACAGTATGCTCTGA 3919
Qy 1861 CAGAACAGTCTGTGGATGACATGATGATGCTTNAAGGNAAGCAGCAGAGAGAGCY 1920
Db 3920 CAGAACAGTCTGTGGATGACATGATGATGCTTAA--GGAAAGCAGCAGAGAGAGAGC- 3976
Qy 1921 TCCGGGCCCCAGNCCCACTNATCAGTGTNCCAGCGTGTCTGTTNCCAGNAGACA 1980
Db 3977 TCCGGGCCCCAGNCCCACTNATCAGTGT-CCAGCGTGTGTTTCCCCAGAGCAGC- 4032
Qy 1981 GCTNCAAGCATCANCACTGACACTNCACTTNGCCCTTGGCCCTTNGGCCANGAGGCTACTG 2040
Db 4033 -----TCAGCATCACTGACACTCACTCCCTGCTGCCCTTGGCCCTTNGGCCANGAGGCTACTG 4083
Qy 2041 CCGNAGCGCATTTTGCACNTCTGATGNACTTCAAGCACTTTTCATGCTNGCCCTTNGG 2100
Db 4084 CCG-ACGGCACTTTTGCACNTCTGATGNACTTCAAGCACTTTTCATGCTNGCCCTTNGG 4140
Qy 2101 GCAGGNCAGGNCAGGNCAGTGCACANTGTAGNAGCATNCCAGGNCAGGNCAGGATGG 2160
Db 4141 GGGCAGGNCAGGNCAGTGCACANTGTAGNAGCATNCCAGGNCAGGNCAGGATGG 4190
Qy 2161 GGTGNAAGGNCAGGNCAGTGTGACGTGTCCANCATGATGTGACTNCCCTCAAACTCTTN 2220
Db 4191 GGTGAA--GGGACNACAGTCTTGTGCTGTCA-CATGATGTGACT-CCTCAAACTCTT- 4245
Qy 2221 NCCAGNATTTCTTAAGATAGCANCCCTTNGCCCACTTTGCCCACTTTAGCTTTCT 2280
Db 4246 --CCAGATTTCTTAAGATAGCANCCCTTNGCCCACTTTGCCCACTTTAGCTTTCT 4301
Qy 2281 CCGAGGNCAGGNCAGTGCACANTGTAGNAGCATNCCAGGNCAGGNCAGGATGG 2340
Db 4302 CCGAGGNCAGGNCAGTGCACANTGTAGNAGCATNCCAGGNCAGGNCAGGATGG 4359

RESULT 10

US-09-981-397A-17

; Sequence 17, Application US/09981397A

; GENERAL INFORMATION:

; APPLICANT: Axxima Pharmaceuticals AG

; APPLICANT: Schubart, Daniel

; APPLICANT: Habenberger, Peter

; APPLICANT: Stein-Gerlach, Matthias

; APPLICANT: Bevec, Dorian

; TITLE OF INVENTION: Cellular Kinases Involved in Cytomegalovirus Infection and their

; TITLE OF INVENTION: Inhibition

; FILE REFERENCE: AXM-004.1 US

; CURRENT APPLICATION NUMBER: US/09/981,397A

; CURRENT FILING DATE: 2002-06-28

; PRIOR APPLICATION NUMBER: 60/240,750

; PRIOR FILING DATE: 2000-10-16

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 17

; LENGTH: 4596

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-981-397A-17

Query Match 70.3%; Score 1850.2; DB 45; Length 4596;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

Qy 1 CCCTCTCACAGCCCGAGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCGCTGT 60
Db 2100 CCCTCTCACAGCCCGAGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCGCTGT 2159
Qy 61 CTGACGCGAGCTGGGAGGAGGTGAACCGGGCACTTACAGCAAGTGGAGGTCTGAAGA 120
Db 2160 CTGACGCGAGCTGGGAGGAGGTGAACCGGGCACTTACAGCAAGTGGAGGTCTGAAGA 2219
Qy 121 GCCTTTGGAGGGAGAAATATAAGAACCAAGACATCCACGCCAAATCAAGCCAAATTACC 180
Db 2220 GCCTTTGGAGGGAGAAATATAAGAACCAAGACATCCACGCCAAATCAAGCCAAATTACC 2279
Qy 181 ACCAGACCTTCCATGCCCGAGAGAGCTTTGCCAGGGCCCGAGGGCCCGGCGAG 240
Db 2280 ACCAGACCTTCCATGCCCGAGAGAGCTTTGCCAGGGCCCGAGGGCCCGGCGAG 2339
Qy 241 CTGAGGAGACAAACGAGCAGAGCCCTTAAGCTCCAGCTCTCTCCACCAAGAGCCCGCAG 300
Db 2340 CTGAGGAGACAAACGAGCAGAGCCCTTAAGCTCCAGCTCTCTCTCCACCAAGAGCCCGCAG 2399
Qy 301 AGCCAAACAAAGTCTCTCTCTTGAAGCAAGAGAGTCTGGGATGTGGAAACCT 360
Db 2400 AGCCAAACAAAGTCTCTCTCTTGAAGCAAGAGAGTCTGGGATGTGGAAACCT 2459
Qy 361 TACTCTGTCTCTCTGGAGCCAGCCCTTCCAGAAACCCAGCTACACAGAGCGGAAG 420
Db 2460 TACTCTGTCTCTCTGGAGCCAGCCCTTCCAGAAACCCAGCTACACAGAGCGGAAG 2519

QY 421 CAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAAATAGAAATTAATTCCTCAACAGCCTGT 480
Db 2520 CAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAAATAGAAATTAATTCCTCAACAGCCTGT 2579
QY 481 CCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGCTCAGCATGCACAGCC 540
Db 2580 CCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGCTCAGCATGCACAGCC 2639
QY 541 TCTCCTCTCGATGACAGTGAGAGAAACCATCAAAAGGCTCTCAAAAGCTCGCGGAGCA 600
Db 2640 TCTCCTCTCGATGACAGTGAGAGAAACCATCAAAAGGCTCTCAAAAGCTCGCGGAGCA 2699
QY 601 CCCTGAGCTCAGGCGTACACTCTTGAGCAGCAGGCGAGGCTCGAAGCTCCAGCTGGA 660
Db 2700 CCCTGAGCTCAGGCGTACACTCTTGAGCAGCAGGCGAGGCTCGAAGCTCCAGCTGGA 2759
QY 661 ACATGGTGTGCGCCGGGGGGGGCCACCGACACCCCAAGCTATTTCATAGGTGGAAG 720
Db 2760 ACATGGTGTGCGCCGGGGGGGGCCACCGACACCCCAAGCTATTTCATAGGTGGAAG 2819
QY 721 TCCAAATACAGTCTTAAATGGTGAAACACTTGACATCCGGAGTTCCACCGGTCAAAG 780
Db 2820 TCCAAATACAGTCTTAAATGGTGAAACACTTGACATCCGGAGTTCCACCGGTCAAAG 2879
QY 781 TGGGAGACATCGCCACTGSCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGCTTGGTCA 840
Db 2880 TGGGAGACATCGCCACTGSCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGCTTGGTCA 2939
QY 841 CCAAGACGGGCAGCGCTTTCGTACGACATGGAGGTGCCAGCTCGGGCATCGACCTGC 900
Db 2940 CCAAGACGGGCAGCGCTTTCGTACGACATGGAGGTGCCAGCTCGGGCATCGACCTGC 2999
QY 901 AGTGACACTGGCCCTGATGGCAGCTTCGCTGAGCTGAGGCTCAAGCATGGCCAGC 960
Db 3000 AGTGACACTGGCCCTGATGGCAGCTTCGCTGAGCTGAGGCTCAAGCATGGCCAGC 3059
QY 961 TGGGAAACAGGCGCTTAAACCTGCTCCACCGCGGCTCCACACTGCGGAAAGCAGCT 1020
Db 3060 TGGGAAACAGGCGCTTAAACCTGCTCCACCGCGGCTCCACACTGCGGAAAGCAGCT 3119
QY 1021 TCTCTCTCGGTGACGATGCTGCTGAAACACAGGCTCAGCCTTCCAGGGGATGTG 1080
Db 3120 TCTCTCTCGGTGACGATGCTGCTGAAACACAGGCTCAGCCTTCCAGGGGATTT - 3177
QY 1081 NCCAGCCCCCGGCTCARGNTGGGAACAGGGCTCGNACGACGAGTGGG 1140
Db 3178 GCCAGCCCCCGGCTCA - CAGTGGGAAACAGGGGCTCG - - - - - CAGCAGCAAGGTGGG 3230
QY 1141 CAAGCNAGATGCCTCCAGGATTTTCAACNCTGAGCCNTGCCCANCCCTGCTGAADA 1200
Db 3231 GCAGCAGATGCTCTCCAGGATTTTCAACNCTGAGCCCTGCCAC - - - - - CCGTCTGA 3284
QY 1201 AAACAYTNCGCCACGCTGAAGAGACAGAGGAGTGGNACGAGGTNNACCTYGGGAA 1260
Db 3285 AAAACATCCGCCACGCTGAAGAGACAGAGGAGGATGGC - - - - - AGGAGTTACCTGGGAA 3340
QY 1261 ACAAACAGGATCTTTNTTCTGCCCTGCTCAGTNCAGTTGGCTGACCTGNACCCGCTTG 1320
Db 3341 ACAAACAGGATCTTTT - TTCTGCCCTGCTCAGT - CGAGTTGGCTGA - - - - - CCGCTG 3396
QY 1321 ANTCTAGTACCATTTTGTGGCAGANCAGGGGAGAGCAGCTTCCAGCCTGGGTGAGAAGG 1380
Db 3397 GATCAGTGACCATTTTGTGGCAGA - CAGGGGAGAGCAGCTTCCAGCCTGGGTGAGAAGG 3455
QY 1381 GTGGGAGCCCTTGGCCCTCACCCTNCCAGGCTGCTGTGNAGAGTCAAGTGTGTA 1440
Db 3456 GTGGGAGCCCTTGGCCCTCACCCT - CCAGGCTGCTGTG - AGAGTGTCAAGTGTGTA 3513
QY 1441 AGGNCCTAAACTCAGTNTTCAGTGCAGAACCAAGTNCAGCAGGTATGCCGCCCGNTA 1500
Db 3514 AGGGCCAAACTC - - - - - AGGTCTAGTGCAGAACAGGT - CAGCAGGTATGCCGCCCGGTAG 3569

QY 1501 GGTTAANNNGGGGCCCTCTNAAACCCCTTCCTCCCTNGGCTNCACCTNGGCCAGCTCANCC 1560
Db 3570 GTTAA - - - - - GGGGGCCTCTAAACCCCTTCCTGGCCTCA - - - - - CCGCCAGCTCA - - - - - 3621
QY 1561 CTTTGGGTAGGGGAAAAGATGCTGACCTTGGGAAGGCTWCCCTGTAGTAATACAC 1620
Db 3622 CTTTGGGTAGGGGAAAAGATGCTGACCTTGGGAAGGCT - CCGTGTAGTAATACAC 3680
QY 1621 CACACTTTTCAGGTTGTGCAACACACAGGTCTCAGTTGACTCTGTGGTTTCAGCCCAAGGACC 1680
Db 3681 CACACTTTTCAGGTTGTGCAACACACAGGTCTCAGTTGACTCTGTGGTTTCAGCCCAAGGACC 3740
QY 1681 AAAGAAGGTGTGAAGTGAAGTGTCTCAGTNCGCCAGACATGTGCCCCCTTTGCTGCTG 1740
Db 3741 AAAGAAGGTGTGAAGTGAAGTGTCTCAGT - CCCCAGACATGTGCCCCCTTTGCTGCTG 3799
QY 1741 GCTACCACTTCTCCCGAGCAGCAGGCCCGGAGCCCTTCAGGCCCAGCAGCTGCCCCAG 1800
Db 3800 GCTACCACTTCTCCCGAGCAGCAGGCCCGGAGCCCTTCAGGCCCAGCAGCTGCCCCAG 3859
QY 1801 ACTCGTGGCACTCAGTTCTCATCTGTAAGGTGAAGGTGATGCAGGATATGCTCTGA 1860
Db 3860 ACTCGTGGCACTCAGTTCTCATCTGTAAGGTGAAGGTGATGCAGGATATGCTCTGA 3919
QY 1861 CAGGAAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGAGCY 1920
Db 3920 CAGGAAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGCAGCAGAGAGAGAGC - 3976
QY 1921 TCCGGCGCCCGAGNCCCACTNATCAGTGTNCCAGCTGTCTGTTNCCCGAGNAGACA 1980
Db 3977 TCCGGCGCCCGAGNCCCACTNATCAGTGT - - - - - CAGCTGTGCTTCCCGAGCAGCAGC - 4032
QY 1981 GCTNCAGNCACTNCACTGACACTNCCCTNGCCCTGCCCCCTNGGCCANGAGGGTACTG 2040
Db 4033 - - - - - TCAGCATCACTGACACTCACCTGCTGCCCTGCCCTGG - - - - - CCAGAGGGTACTG 4083
QY 2041 CCGNACGGCACTTTGCACTGATGACCTTCAAGACACTTTTATGGGTNGCCCTCTNNG 2100
Db 4084 CCG - ACGGCACTTTGCACT - - - - - CTGATGACCTTCAAGCACTTTTATGGGTGCTGCTGCA 4140
QY 2101 GCAGGNCAGGNCAGGNCAGTGCACANCTGTAGNAGCATANGCAAGCCAGAGAGTGG 2160
Db 4141 GGGCAGGNCAGGNCAGTGCACACT - - - - - GTAGAGCATAGCAAGCCAGAGATGG 4190
QY 2161 GGTGNAAGGGANCAAGTCTTGAAGTGTCCANCACTGATGTGACTNCTCAAACTCTTN 2220
Db 4191 GGTGNA - - - - - GGACACAGCTTGTGAGTGTCCA - CATGATGTGACT - CCTCAAACTCTT - 4245
QY 2221 NCCAGNATTTCTTAAAGATAGANCCCTTNCCTTCCCATGCCCCAGCTTAGCCTCTTCT 2280
Db 4246 - - - - - CCAGATTTCTTAAAGATAGCACCCTT - - - - - TTCCCATTTGCCCGAGCTTAGCCTCTTCT 4301
QY 2281 CCCAGGGGAGCTNCTCAGGACTCAGCTAGCATTAATACAGCTGTGNAATCGTCAGGGGG 2340
Db 4302 CCCAGGGGAGCTA - CTAGGACTCAGCTAGCATTAATACAGCTGTG - AATGTCAGGGGG 4359
QY 2341 TGTCTGTAGCTCAACCTCTCTGGGGCAGGGGAGCGCCAGAGACTCCGTGGGAGAAAGCTCAT 2400
Db 4360 TGTCTGTAGCTCAACCTCTCTGGGGCAGGGGAGCGCCAGAGACTCCGTGGGAGAAAGCTCAT 4419
QY 2401 TCCACATCTTTCGCAAGACAGCTTTTNGTCAAGCTGTCCATGATGAGTCAGAGCTCTCC 2460
Db 4420 TCCACATCTTTCGCAAGACAGCTTTT - GTCCAGCTGTCCCATTTGAGTCAGAGCTCTCC 4478
QY 2461 GGGGAGAGAGCCCGGGCCCCCAGACATAAAGAACTGCAGCTTGGTACTGCAGAGTCTG 2520
Db 4479 GGGGAGAGAGCCCGGGCCCCCAGACATAAAGAACTGCAGCTTGGTACTGCAGAGTCTG 4538
QY 2521 GGTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGATGACAAATGTTAAAAA 2578
Db 4539 GGTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGATGACAAATGTTAAAAA 4596

RESULT 11

US-10-172-118-975
; Sequence 975, Application US/10172118
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 975
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NW_003954
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-975

Query Match 70.3%; Score 1850.2; DB 49; Length 4596;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

Qy 1 CCCCTCTCACAGCCAGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
Db 2100 CCCCTCTCACAGCCAGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 2159

Qy 61 CTGACGCGAGCTGGAGGGAAGTGAAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120
Db 2160 CTGACGCGAGCTGGAGGGAAGTGAAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 2219

Qy 121 GCCCTTGGAGGGGAGATATAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 180
Db 2220 GCCCTTGGAGGGGAGATATAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 2279

Qy 181 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCGAGGGCCCGGCCAG 240
Db 2280 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCGAGGGCCCGGCCAG 2339

Qy 241 CTGAGGAGACAAAGGCGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCGCCAG 300
Db 2340 CTGAGGAGACAAAGGCGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCGCCAG 2399

Qy 301 AGCCAAACAACTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGAGGAAACCT 360
Db 2400 AGCCAAACAACTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGAGGAAACCT 2459

Qy 361 TACCT 420
Db 2460 TACCT 2519

Qy 421 CAACCGTCCCGAGCAGGAATCGCAGCAGCTGGAAATAGAAATATTCTCTCAACAGCCTGT 480
Db 2520 CAACCGTCCCGAGCAGGAATCGCAGCAGCTGGAAATAGAAATATTCTCTCAACAGCCTGT 2579

Qy 481 CCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATCTCTCTGCTCAGCATGACAGCC 540
Db 2580 CCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATCTCTCTGCTCAGCATGACAGCC 2639

Qy 541 TCTCCCTGTCGATGACAGTACAGAGACCCATCAAGGGCTCTCAAGCTCGCGGACA 600
Db 2640 TCTCCCTGTCGATGACAGTACAGAGACCCATCAAGGGCTCTCAAGCTCGCGGACA 2699

Qy 601 CCCTGAGCTCAGGGGTACACTCTCTGGAGCAGCCAGGGCTCGAAGCTCCAGCTGGA 660

Db 2700 CCCTGAGCTCAGGGGTACACTCTTGAGAGCAGCCAGGGCTCGAAGCTCCAGCTGA 2759
Qy 661 ACATGGTCTGCGCCGGGGGGCCACCCGACACCCCAAGCTATTTCATGGTGTGAAG 720
Db 2760 ACATGGTCTGCGCCGGGGGGGGCCACCCGACACCCCAAGCTATTTCATGGTGTGAAG 2819
Qy 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGSCACATCCGGGAGTTCACCGGGTCAAAG 780
Db 2820 TCCAAATACAGTCTCTTAATGGTGAACACCTGSCACATCCGGGAGTTCACCGGGTCAAAG 2879
Qy 781 TGGGAGACATCGCCACTGGGATCAGCAGCAGATCCAGCTGCAGCCTTCAGCTTGGTCA 840
Db 2880 TGGGAGACATCGCCACTGGGATCAGCAGCAGATCCAGCTGCAGCCTTCAGCTTGGTCA 2939
Qy 841 CCAAAGACGGCAGCCTGTTCTGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC 900
Db 2940 CCAAAGACGGCAGCCTGTTCTGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC 2999
Qy 901 AGTGACACATGGCCCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGC 960
Db 3000 AGTGACACATGGCCCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGC 3059
Qy 961 TGGAGAACAGGCCCTTAACCTGCCCCCTCACCGCGGCTCCACACTGCCGGAAGAGAGCCT 1020
Db 3060 TGGAGAACAGGCCCTTAACCTGCCCCCTCACCGCGGCTCCACACTGCCGGAAGAGAGCCT 3119
Qy 1021 TCCTGCTCGGTGACGATGCTGCCCTGAAACACAGAGCTGCAGCTTCCAGGGGATTTG 1080
Db 3120 TCCTGCTCGGTGACGATGCTGCCCTGAAACACAGAGCTGCAGCTTCCAGGGGATTTG 3177
Qy 1081 NCCAGCCCCCGGCTCARCAGNTGGGAACACAGGGCTTCGNCAGCAGNAGCNAAGTNGGGGG 1140
Db 3178 GCAGACCCCCCGGCTCA--CAGTGGGAACACAGGGCTCG-----CAGCAGCAAGTGGGG 3230
Qy 1141 CAAGCNAAGATGCTCCAGGATTTTCAACCTGAGCCCTGAGCCCTGAGCCCTGCTGAADA 1200
Db 3231 GCAAGCAAGATGCTCCAGGATTTTCAACCTGAGCCCTGAGCCCTGAGCCCTGCTGAADA 3284
Qy 1201 AAACAYTNCGCCACGTGAAGAGACAGAGAGATGNCAGAGGATTTNNACTTGGGGAA 1260
Db 3285 AAAAACATCGCCACGTGAAGAGACAGAGAGATGNCAGAGGATGNC-----AGGAGTTACCTGGGGAA 3340
Qy 1261 ACAAAACAGGGATCTTTTCTGCCCCCTGCTCCAGTNCAGCTTGGCCCTGNAACCCGCTTGG 1320
Db 3341 ACAAAACAGGGATCTTTTCTGCCCCCTGCTCCAGT--CGAGTGGCCCTGA--CCCGCTTG 3396
Qy 1321 ANTCAGTGACCATTTTGTGGCAGANCAGGGGAGAGAGCTTCCAGCCTGGGTCAAGAGGG 1380
Db 3397 GATCAGTGAACCATTTTGTGGCAGA--CAGGGGAGAGAGCTTCCAGCCTGGGTCAAGAGGG 3455
Qy 1381 GTGGGCGAGCCTTCGGGCCCTCACCCCTNCAGGCTGCTGTGNAGAGTGTCAAGTGTGTA 1440
Db 3456 GTGGGCGAGCCTTCGGGCCCTCACCCCT--CCAGGCTGCTGTG--AGAGTGTCAAGTGTGTA 3513
Qy 1441 AGGNCNCCAAANTCAGGNTTTCAGTGCAAGAACAGGTNACAGAGTATGCCCGCCCGNTA 1500
Db 3514 AGGNCNCCAAACTC---AGGTTTCAGTGCAAGAACAGGT--CAGCAGGATATGCCCGCCCGTAG 3569
Qy 1501 GGTTAANNNGGGGGCCCTCTNAAACCCCTTGCCTNGGCCCTNACCTNGSCCAGCTCANCCC 1560
Db 3570 GTTAA----GGGGGCCCTCTAAACCCCTTGCCTGGCCCTCA--CCTGGCAGCTCA--CCC 3621
Qy 1561 CTTTTGGGTGTAGGGGAAAAGATGCTGACCCCTGGGAAGGCTGCCCTGGTAGAATACAC 1620
Db 3622 CTTTTGGGTGTAGGGGAAAAGATGCTGACCCCTGGGAAGGCT--CCCTGGTAGAATACAC 3680
Qy 1621 CACATTTTTCAGGTGTGTCAAACACAGGTCTGAGTGTGAACCTCTGCTGTTTCAAGGAGACC 1680
Db 3681 CACATTTTTCAGGTGTGTCAAACACAGGTCTGAGTGTGAACCTCTGCTGTTTCAAGGAGACC 3740
Qy 1681 AAAGAAGGTGTGAAGTGAAGTGTCTTCAGTNC--CCAGACATGTGCCCTTTGCTGCTG 1740

Db 3741 AAAGAAGTGTGAAGTGAAGTGTCTCAGT-CCCAGACATGTGCCCTTTGCTGCTG 3799
Qy 1741 GCTACCACTTTCCCGAGAGCAGAGGCCCCGAGCCCTTCAGGCCAGCACTGCCCGAG 1800
Db 3800 GCTACCACTTTCCCGAGAGCAGAGGCCCCGAGCCCTTCAGGCCAGCACTGCCCGAG 3859
Qy 1801 ACTGCTGGCACTGATGCTCCCTCATCTCTAAGGTGAAGGTGATGAGGATATGCTGA 1860
Db 3860 ACTGCTGGCACTGATGCTCCCTCATCTCTAAGGTGAAGGTGATGAGGATATGCTGA 3919
Qy 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAAGGNAAGCAGCAGAGAGAGCGY 1920
Db 3920 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAAGGNAAGCAGCAGAGAGAGCG- 3976
Qy 1921 TCCGCGCCCCAGNCCCACACTNATCAGTGTGTCAGGCTGCTGCTGCTGCTGCTGCTG 1980
Db 3977 TCCGCGCCCCAGNCCCACACTNATCAGTGTGTCAGGCTGCTGCTGCTGCTGCTGCTG 4032
Qy 1981 GCTNCAGNATCANCACACTGACATCNCACCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
Db 4033 -----TCAGATCACACTGACATCNCACCTGCTGCTGCTGCTGCTGCTGCTGCTG 4083
Qy 2041 CCGNACGGCACTTTGCACTGATGATGATGATGATGATGATGATGATGATGATGATG 2100
Db 4084 CCG-ACGGCACTTTGCACTGATGATGATGATGATGATGATGATGATGATGATGATG 4140
Qy 2101 GCAGGNCAGGNCAGGNCAGTGTGATGATGATGATGATGATGATGATGATGATGATG 2160
Db 4141 GGGCAGGGCAGGGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 4190
Qy 2161 GGTGNAAGGNCACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2220
Db 4191 GGTGAA--GGGACACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 4245
Qy 2221 NCCAGNATTTCTTAAGATAGCAGTGTGATGATGATGATGATGATGATGATGATGAT 2280
Db 4246 --CGATTTCTTAAGATAGCAGTGTGATGATGATGATGATGATGATGATGATGATG 4301
Qy 2281 CCGAGGGAGTANCTCAGGACTCAGTGTGATGATGATGATGATGATGATGATGATGATG 2340
Db 4302 CCGAGGGAGTANCTCAGGACTCAGTGTGATGATGATGATGATGATGATGATGATGATG 4359
Qy 2341 TGTCTGCTAGCTCAACCTCTTGGGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
Db 4360 TGTCTGCTAGCTCAACCTCTTGGGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4419
Qy 2401 TCCACATCTTGCCAAAGACAGCTTTGTCAGGCTGTCAGGCTGTCAGGCTGTCAGGCT 2460
Db 4420 TCCACATCTTGCCAAAGACAGCTTTGTCAGGCTGTCAGGCTGTCAGGCTGTCAGGCT 4478
Qy 2461 GGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
Db 4479 GGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4538
Qy 2521 GGTGTGAGAACTCTTTGTAAGCAATAAAGTTTGGGGTGTGATGACAAATGTTAAAAA 2578
Db 4539 GGTGTGAGAACTCTTTGTAAGCAATAAAGTTTGGGGTGTGATGACAAATGTTAAAAA 4596

RESULT 12
US-10-283-975A-294
; Sequence 294, Application US/10283975A
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
; FILE REFERENCE: CDS 293 PCT
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/340,938
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/338,997
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/340,081

; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/341,012
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 294
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: HUMAN
; US-10-283-975A-294

Query Match 70.3%; Score 1850.2; DB 51; Length 4596;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;
Qy 1 CCGCTCTCACAGCCCGAGCCATCCAAAGAGGGGCTGAGGAAGAGCCCATCCACCGGCTGT 60
Db 2100 CCGCTCTCACAGCCCGAGCCATCCAAAGAGGGGCTGAGGAAGAGCCCATCCACCGGCTGT 2159
Qy 61 CTGAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTTACAGCAAGTGGGAGGCTTGAAGA 120
Db 2160 CTGAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTTACAGCAAGTGGGAGGCTTGAAGA 2219
Qy 121 GCGCTTGGAGGGAGATATTAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC 180
Db 2220 GCGCTTGGAGGGAGATATTAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC 2279
Qy 181 ACCAGACCTTCCATGCCAGCCAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCGCGCAG 240
Db 2280 ACCAGACCTTCCATGCCAGCCAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCGCGCAG 2339
Qy 241 CTGAGGAGACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
Db 2340 CTGAGGAGACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2399
Qy 301 AGCCAAACAAGTCTCCTCCCTTGAAGTCTGAGCAAGGAGGAGTCTGGATGTGGAAACCT 360
Db 2400 AGCCAAACAAGTCTCCTCCCTTGAAGTCTGAGCAAGGAGGAGTCTGGATGTGGAAACCT 2459
Qy 361 TACTCTGCTCCTCCTGAGCCAGCCCTTGCAGAAACCCAGCTCACCAGAGCGGAAAG 420
Db 2460 TACTCTGCTCCTCCTGAGCCAGCCCTTGCAGAAACCCAGCTCACCAGAGCGGAAAG 2519
Qy 421 CAACCGTCCCGAGAGGAGAACTGCAGAGCTGGAAATAGATTTTCTCAACAGCCCTGT 480
Db 2520 CAACCGTCCCGAGAGGAGAACTGCAGAGCTGGAAATAGATTTTCTCAACAGCCCTGT 2579
Qy 481 CCCAGCAATTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTCTCAGCATCGACAGCC 540
Db 2580 CCCAGCAATTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTCTCAGCATCGACAGCC 2639
Qy 541 TCTCCTCTCGGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGGCTCGCGGAGCA 600
Db 2640 TCTCCTCTCGGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGGCTCGCGGAGCA 2699
Qy 601 CCGTCTGAGTCTGAGCTGAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
Db 2700 CCGTCTGAGTCTGAGCTGAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2759
Qy 661 ACATGCTGCTGCGCGGG 720
Db 2760 ACATGCTGCTGCGCGGG 2819
Qy 721 TCCAAATACAGTCTCTTAATGTTGAACACCTGACATCCCGGAGTTCACCGGGTCAAG 780
Db 2820 TCCAAATACAGTCTCTTAATGTTGAACACCTGACATCCCGGAGTTCACCGGGTCAAG 2879
Qy 781 TGGGAGACATCGCCACTGTCATCAGCAGCAGATCCAGCTCAGCTTCAAGCTTCAAGCTTCA 840
Db 2880 TGGGAGACATCGCCACTGTCATCAGCAGCAGATCCAGCTCAGCTTCAAGCTTCAAGCTTCA 2939
Qy 841 CCAAGAGCGGGGAGGCTGTTCTGCTACGACATGAGGTGCAGATCGGGGCTCGAGCTGCG 900

Db 2940 CCAAGACGGGACGCTGTTGCTACGACATGAGGTGCCAGACTCGGGCATGACCTGC 2999
Qy 901 AGTGCACTGGCCCTGATGGAGCTTGGCTGGAGCTGAGGTCAAGCATGGCCAGC 960
Db 3000 AGTGCACTGGCCCTGATGGAGCTTGGCTGGAGCTGAGGTCAAGCATGGCCAGC 3059
Qy 961 TGGAGACAGGCGCTTACCCCTGACCTCCACCGCGGCTCCACATGCGGAAAGCAGCT 1020
Db 3060 TGGAGAACAGGCGCTTAAACCTGACCTCCACCGCGGCTCCACATGCGGAAAGCAGCT 3119
Qy 1021 TCCTGCTGGTGACGATGCTGCTGAAACACAGAGCTCAGCGCTTCCAGGGGATVTG 1080
Db 3120 TCCTGCTGGTGACGATGCTGCTGAAACACAGAGCTCAGCGCTTCCAGGGGATVTG 3177
Qy 1081 NCGAGCCCCCGCTCARGAGTGGGAAACAGGCTCGCAGGCTGCGNAGNAGNAGGTTGGGG 1140
Db 3178 GCCAGCCCCCGCTCA--CAGTGGGAACAGGCGCTCG-----CAGCAGCAAGGTGGG 3230
Qy 1141 CAAGNAGATGCTCCCGAGATTTCCANCTGAGCCCTGCTGCTGAADA 1200
Db 3231 GCAAGCAGATGCTCCCGAGATTTCCACCTGAGCCCTGCTGCTGA 3284
Qy 1201 AAAAYTNCCCGCAGTGAAGAGACAGAAAGAGGATGNCAGGAGTTNACCTTGGGGAA 1260
Db 3285 AAAACATCCCGCAGTGAAGAGACAGAAAGAGATGCG-----AGGAGTTACCTGGGGAA 3340
Qy 1261 ACAAAAAGGAGATTTTNTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db 3341 ACAAAAAGGAGATTTT--TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3396
Qy 1321 ANTCACTGACATTTGCTGAGAGAGAGGAGAGAGCTTCCAGCTGGGTCAAGAGGG 1380
Db 3397 GATCAGTGACATTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3455
Qy 1381 GTGGGAGAGCCTTGGGCGCTTCACTTCCAGGCTGCTGAGAGAGAGAGAGAGAGAGAG 1440
Db 3456 GTGGGAGAGCCTTGGGCGCTTCACTTCCAGGCTGCTGAGAGAGAGAGAGAGAGAG 3513
Qy 1441 AGGNCCTAAANCTCAGTNTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 3514 AGGNCCTAAANCTCAGTNTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3569
Qy 1501 GGTAAANNNGGGGCTTAAACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
Db 3570 GTTAA---GGGGGCTTCTAAACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3621
Qy 1561 CTTTGGGTGAGGGAAG 1620
Db 3622 CTTTGGGTGAGGGAAG 3680
Qy 1621 CACACTTTTTCAGGTTGTTGCAACAGAGTCTGAGTTGACCTCTGTTGTTGAGCCAGGAC 1680
Db 3681 CACACTTTTTCAGGTTGTTGCAACAGAGTCTGAGTTGACCTCTGTTGTTGAGCCAGGAC 3740
Qy 1681 AAGAAGGTGTAAGTGAAGTGTCTGAGTNTCCAGAGATGTCCTTCTGCTGCTG 1740
Db 3741 AAGAAGGTGTAAGTGAAGTGTCTGAGTNTCCAGAGATGTCCTTCTGCTGCTG 3799
Qy 1741 GCTACCACTTCTCCAG 1800
Db 3800 GCTACCACTTCTCCAG 3859
Qy 1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGTGAAGGTGATGAGGATATGCTGTA 1860
Db 3860 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGTGAAGGTGATGAGGATATGCTGTA 3919
Qy 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTGAGTGTCTGAGGAGAGAGAGAG 1920
Db 3920 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTGAGTGTCTGAGGAGAGAGAGAG 3976
Qy 1921 TCGGCGCGCCAGNCCCACTNATCAGTGTNCCAGCGTGTGCTGCTGCTGCTGCTGCTG 1980
Db 3977 TCGGCGCGCCAGNCCCACTNATCAGTGTNCCAGCGTGTGCTGCTGCTGCTGCTGCTG 4032

Qy 1981 GCTNCAAGCATCANCACTGACACTNCACTNCACTNCACTNCACTNCACTNCACTNCACT 2040
Db 4033 -----TCAGCATCACTGACACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 4083
Qy 2041 CCGNACCGCACTTTCGACNTCTGATGACCTCAAGCACTTTCATGCTGCTGCTGCTGCTGCT 2100
Db 4084 CCG-ACGGCACTTTCGACNTCTGATGACCTCAAGCACTTTCATGCTGCTGCTGCTGCTG 4140
Qy 2101 GCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAG 2160
Db 4141 GGGCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAG 4190
Qy 2161 GGTNAAAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGG 2220
Db 4191 GGTGAA--GGGACACAGTCTTGAAGTGTGCA--CATGATGTGACT--CTCAAAACCTCTT- 4245
Qy 2221 NCCAGNATTTCTTAAAGATAGCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAG 2280
Db 4246 --CCAGATTTCTTAAAGATAGCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAG 4301
Qy 2281 CCCAGGAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGG 2340
Db 4302 CCCAGGAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGG 4359
Qy 2341 TGCTGCTAGCCTCAACCTCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
Db 4360 TGCTGCTAGCCTCAACCTCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4419
Qy 2401 TCCACATCTTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2460
Db 4420 TCCACATCTTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4478
Qy 2461 GGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
Db 4479 GGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4538
Qy 2521 GGTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2578
Db 4539 GGTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4596

RESULT 13

US-10-342-887-975
; Sequence 975, Application US/10342887
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 975
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-975

Query Match 70.3%; Score 1850.2; DB 52; Length 4596;
Best Local Similarity 90.5%; Pred. No. 0;

Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

QY	1	CCCCTCTCACAGCCACAGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCGTGT	60	1081	NCCAGCCCCCGGCTCARCAGNTGGGAACCAAGGCCTTCGNCAGCNAGCNAAAGTNGGGG	1140
Db	2100	CCCCCTCTCACAGCCACAGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCGTGT	2159	3178	GCCAGCCCCCGGCTCA--CAGTGGGAACCAAGGCCTCG-----CAGCAGCAAGGTGGG	3230
QY	61	CTGCAGCGAGCTGGGAGGAGGTGAACCCGGGCACTACAGCAAGTGGGAGGTCTGAAGA	120	1141	CAAGCNAGATGCTTCCAGGATTTTCANAGCTTTCAGCCCTGAGCCCTGAGCCCTGCTGAADA	1200
Db	2160	CTGCAGCGAGCTGGGAGGAGGTGAACCCGGGCACTACAGCAAGTGGGAGGTCTGAAGA	2219	3231	GCAAGCAGATGCTTCCAGGATTTTCACCTGAGCCCTGCCCCAC-----CCTGTCTGA	3284
QY	121	GCCCTTGAGGGGAGAAATATAAGAACCAAGACATCCACCGCAATCAAGCCAAATTACC	180	1201	AAACAYTNCGCCACAGTGAAGAGACAGAGAGATGNCAGGAGTTNNACTTGGGAA	1260
Db	2220	GCCCTTGAGGGGAGAAATATAAGAACCAAGACATCCACCGCAATCAAGCCAAATTACC	2279	3285	AAAAACATCCCCACAGTGAAGAGACAGAGAGATGGC-----AGGAGTTACTTGGGAA	3340
QY	181	ACCAGACCTTCATSCCAGCCGAGAGCTTTGSCCAAGGCCCCAGGGCCCGGCCAG	240	1261	ACAAAAAGGAGATCTTTTCTGCCCCCTGCTCCAGTNCAGTTCGCTGNAACCCGCTTGG	1320
Db	2280	ACCAGACCTTCATSCCAGCCGAGAGCTTTGSCCAAGGCCCCAGGGCCCGGCCAG	2339	3341	ACAAAAAGGAGATCTTTTCTGCCCCCTGCTCCAGT--CGAGTTGGCCTGA--CCGCTTG	3396
QY	241	CTGAGGAGACAACAGGCAGAGCCCTTAAGCTCCAGCCTCCTCTCCACAGAGCCCCAG	300	1321	ANTCAGTGACCATTTGTTGGCAGANACAGGGGAGAGAGCTTTCAGCTTGGGTCAAGGG	1380
Db	2340	CTGAGGAGACAACAGGCAGAGCCCTTAAGCTCCAGCCTCCTCTCCACAGAGCCCCAG	2399	3397	GATCAGTGACCATTTGTTGGCAGA--CAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGGG	3455
QY	301	AGCCAAAACAAGTCTCCTCCTTGAATTTGACGACGAGGAGTCTGGGATGTGGGAACCT	360	1381	GTGGCGAGCCCTTCGGCCCTCACCCTNCCAGGCTGCTGTGNAGAGTGTCAAGTGTGTA	1440
Db	2400	AGCCAAAACAAGTCTCCTCCTTGAATTTGACGACGAGGAGTCTGGGATGTGGGAACCT	2459	3456	GTGGCGAGCCCTTCGGCCCTCACCCT--CAGGCTGCTGTG--AGAGTGTCAAGTGTGTA	3513
QY	361	TACCTCTGTCTCTGGAGCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAG	420	1441	AGGNCCCAANCTCAGNTTTCAGTGCAGAACCCAGGTNCCAGAGTATGCCGCCCGGNTA	1500
Db	2460	TACCTCTGTCTCTGGAGCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAG	2519	3514	AGGCCCAAACTC--AGTTCAGTGCAGAACCAAGGT--CAGCAGTATGCCGCCCGTAG	3569
QY	421	CAACCGTCCCGAGCAGGAATGTCAGCAGCTGGAATAGAAATATTCCTCAACAGCCTGT	480	1501	GTTTAA--G--GGGGCCCTCTAAACCCCTTGCTGCGCTCA---CCTGGCCAGCTCA--CCC	3621
Db	2520	CAACCGTCCCGAGCAGGAATGTCAGCAGCTGGAATAGAAATATTCCTCAACAGCCTGT	2579	1561	CTTTTGGGTGAGGGGAAAGAAATGCTGAGCCCTGGGAAGGCTWCCTCTGGTAGAATACAC	1620
QY	481	CCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTCCTGTGCTCAGCATCGACGCC	540	3622	CTTTTGGGTGAGGGGAAAGAAATGCTGAGCCCTGGGAAGGCT--CCCTGGTAGAATACAC	3680
Db	2580	CCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTCCTGTGCTCAGCATCGACGCC	2639	1621	CACACTTTTCAGGTTGTTGCAACACAGTCTGAGTTGACCTCTGTTTCAGCCAGGAGCC	1680
QY	541	TCCTCCCTGTCGATCAGTGAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGGACA	600	3681	CACACTTTTCAGGTTGTTGCAACACAGTCTGAGTTGACCTCTGTTTCAGCCAGGAGCC	3740
Db	2640	TCCTCCCTGTCGATCAGTGAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGGACA	2699	1681	AAAGAGGTGTAAAGTAAAGTGTTCAGTNCNCCAGACATGTGCCCTTGTGCTGTG	1740
QY	601	CCCTGAGCTCAGCGCTACACTCTCTGAGCAGCAGGCGGAGCTCGAAGCTCCAGCTGA	660	3741	AAAGAGGTGTAAAGTAAAGTGTTCAGT--CCCCAGACATGTGCCCTTGTGCTGTG	3799
Db	2700	CCCTGAGCTCAGCGCTACACTCTCTGAGCAGCAGGCGGAGCTCGAAGCTCCAGCTGA	2759	1741	GCTACCACTCTTCCCAGAGCAGGCGCCGAGCCCTTCAGGCCCAGCACTGCCCCAG	1800
QY	661	ACATGGTGTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGCTGGAAG	720	3800	GCTACCACTCTTCCCAGAGCAGGCGCCGAGCCCTTCAGGCCCAGCACTGCCCCAG	3859
Db	2760	ACATGGTGTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGCTGGAAG	2819	1801	ACTCGTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTGTATGCAGGATATGCCCTGA	1860
QY	721	TCCAAATACAGTCTCTTAATGTTGAACACCTGCACATCGGAGTTCCACCGGGTCAAAG	780	3860	ACTCGTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTGTATGCAGGATATGCCCTGA	3919
Db	2820	TCCAAATACAGTCTCTTAATGTTGAACACCTGCACATCGGAGTTCCACCGGGTCAAAG	2879	1861	CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAGGNAAGACAGCAGAGAGAGCY	1920
QY	781	TGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCAGCTGCAGCTTCAGCTTGGTCA	840	3920	CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAGGNAAGACAGCAGAGAGAGCG-	3976
Db	2880	TGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCAGCTGCAGCTTCAGCTTGGTCA	2939	1921	TCCGCGCCCCCAGNCCCACCTNATCAGTGTNCCAGCGTGTNGTTCNCCCAGNAGACACA	1980
QY	841	CCAAAGACGGGAGCCTGTTGCTAGCAATGGAGGTGCCAGATCGGGCATCGACCTGC	900	3977	TCCGCGCCCCCAGNCCCACCTNATCAGTGT--CCAGCTGTCTGTTTCCCAGNAGCAGC-	4032
Db	2940	CCAAAGACGGGAGCCTGTTGCTAGCAATGGAGGTGCCAGATCGGGCATCGACCTGC	2999	1981	GCTNCAGNCACTCANCACTGACACTNCACTNCGCTTCCCTCCCTNCGCCAGNAGGTA	2040
QY	901	AGTGACACTGCGCCCTGATGACGCTTCGCTCGAGCTGGAGGTCAAGCATGCGCAGC	960	4033	-----TCAGCATCACACTGACACTCACCTCCCTGCTCCCTGCTG--CCAGAGGGTACTG	4083
Db	3000	AGTGACACTGCGCCCTGATGGAGCTTCGCTCGAGCTGGAGGTCAAGCATGCGCAGC	3059	2041	CCGNAAGCAGCTTTCACATCTGATGNAACCTCAAGCAGCTTTCATGGCTGCTCTNNG	2100
QY	961	TGGAGAACAGGCCCTAACCTGCCCTTCCACCGCGCTCCACACTGCCGGAAGCAGCT	1020	4084	CCG--ACGCGCACTTTCGACT--CTGATGACCTCAAGCAGCTTTCATGGCTGCTCTGGA	4140
Db	3060	TGGAGAACAGGCCCTAACCTGCCCTTCCACCGCGCTCCACACTGCCGGAAGCAGCT	3119	2101	GCAGGNCAGGNCAGGNCAGTGCANCTGTAGNAGCATANGCAANGCCAGGAGATGG	2160
QY	1021	TCTGTCTGCTGACATGCTGCTGGAACACAGGCTCAGCGCTTCCAGGGGATTTG	1080	4141	GGCAGGCGCAGGCGAGTGACACT-----GTAGGAGCATACCAAGCCAGGAGATGG	4190
Db	3120	TCTGTCTGCTGACATGCTGCTGGAACACAGGCTCAGCGCTTCCAGGGGATTT	3177			

Db	3456	GTGGGAGCCCTTCGGCCCTCAACCTT-CCAGCTGCTGTG-AGAGTGTCAAGTGTGTA	3513		Db	4479	GGGGAGAGAGCCCGCCCGCCAGACATAAAGAACTGCAGCCCTTGGTACTGCAGAGTCTG	4538
Qy	1441	AGGNCACAACTCAGTTCAGTGCAGAACAGGTCAGCAGGATGCCGCCCGGNTA	1500		Qy	2521	GTTTGTAGAGAACTCTTTGTAAAGCAATAAGTTTGGGGTGATGACAAATGTTAAAAA	2578
Db	3514	AGGGCCCAACTC---AGGTTTCACTGAGAGAACCAAGGT-CAGCAGGTATGCCGCCCGGTAG	3569		Db	4539	GTTTGTAGAGAACTCTTTGTAAAGCAATAAGTTTGGGGTGATGACAAATGTTAAAAA	4596
Qy	1501	GTTTAAANGGGGCGCTCTNAAACCCCTTGCTTNGGCTCNCACCTNGGCCAGCTTCANCCC	1560		RESULT 15			
Db	3570	GTTAA---GGGGGCGCTTAAACCCCTTGCTTGCCTCA---CCTGCCAGCTCA-CCC	3621		US-09-496-914A-6881			
Qy	1561	CTTTTGGGTAGGGGAAAGAAATCCCTGACCCCTGGGAAGCTCCCTGGTGTAGAAATAC	1620		; Sequence 6881, Application US/09496914A			
Db	3622	CTTTTGGGTAGGGGAAAGAAATCCCTGACCCCTGGGAAGCT-CCCTGGTGTAGAAATAC	3680		; GENERAL INFORMATION:			
Qy	1621	CACACTTTTCACTGTTGCAACACAGTCTCTGAGTTCAGCTCTGTTTCAGCCAGGACC	1680		; APPLICANT: Tang, Yuanhua T.			
Db	3681	CACACTTTTCACTGTTGCAACACAGTCTCTGAGTTCAGCTCTGTTTCAGCCAGGACC	3740		; APPLICANT: Tillinghast, John			
Qy	1681	AAAGAAGGTGTGTAAAGTGAAGTGTCTCAGTNCACAGATGTCGCCCTTTGCTGTG	1740		; APPLICANT: Sinku, Ankura			
Db	3741	AAAGAAGGTGTGTAAAGTGAAGTGTCTCAGT-CCCCAGACATGTGCCCTTTGCTGTG	3799		; APPLICANT: Liu, Chenghua			
Qy	1741	GCTACACTCTTCCAGAGCAGCAGGCCGCCAGCCCTTCAAGGCCAGCACTGCCCCAG	1800		; APPLICANT: Drmanac, Radoje T.			
Db	3800	GCTACACTCTTCCAGAGCAGCAGGCCGCCAGCCCTTCAAGGCCAGCACTGCCCCAG	3859		; TITLE OF INVENTION: Novel Contigs Obtained			
Qy	1801	ACTCGTGGCACTCAGTTCCTTCACTGTAAAGTGAAGGTGATGACAGGATATGCTGA	1860		; TITLE OF INVENTION: From Various Libraries			
Db	3860	ACTCGTGGCACTCAGTTCCTTCACTGTAAAGTGAAGGTGATGACAGGATATGCTGA	3919		; FILE REFERENCE: 787			
Qy	1861	CAGGAACAGTCTGTGTGATGGAATGATCAGTCTGTAAGNAAAGCAGCAGAGAGAGY	1920		; CURRENT APPLICATION NUMBER: US/09/496,914A			
Db	3920	CAGGAACAGTCTGTGTGATGGAATGATCAGTCTGTAAGNAAAGCAGCAGAGAGAGY	3976		; CURRENT FILING DATE: 2000-02-03			
Qy	1921	TCCGGGCCCCAGNCCCCTACTCAGTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1980		; PRIOR APPLICATION NUMBER: US 09/353,690			
Db	3977	TCCGGGCCCCAGNCCCCTACTCAGTTCAGTGT---CCAGGCTGCTGTTCCCGCAGCAGC-	4032		; PRIOR FILING DATE: 1999-07-14			
Qy	1981	GCTNCAGNATCANCACTGACACTNCACTGACACTGACACTGACACTGACACTGACACTG	2040		; PRIOR APPLICATION NUMBER: US 09/034,341			
Db	4033	-----TCAGATCACACTGACACTGACACTGACACTGACACTGACACTGACACTG	4083		; PRIOR FILING DATE: 1998-02-13			
Qy	2041	CCGACGCACTTGTGACACTGACACTGACACTGACACTGACACTGACACTGACACTG	2100		; PRIOR APPLICATION NUMBER: US 09/045,400			
Db	4084	CCG-ACGCACTTGTGACACT--CTGATGACCTCAAGCACTTTCATGGCTGCCCTCTGCA	4140		; PRIOR FILING DATE: 1998-03-20			
Qy	2101	GCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAG	2160		; PRIOR APPLICATION NUMBER: US 09/321,214			
Db	4141	GGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	4190		; PRIOR FILING DATE: 1999-05-26			
Qy	2161	GGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2220		; PRIOR APPLICATION NUMBER: US 09/131,598			
Db	4191	GGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	4245		; PRIOR FILING DATE: 1998-08-10			
Qy	2221	NCCAGNATTTCTTCAAGATAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2280		; PRIOR APPLICATION NUMBER: US 09/431,517			
Db	4246	--CCAGATTTCTTCAAGATAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	4301		; PRIOR FILING DATE: 1999-11-01			
Qy	2281	CCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2340		; PRIOR APPLICATION NUMBER: US 09/328,351			
Db	4302	CCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	4359		; PRIOR FILING DATE: 1999-06-04			
Qy	2341	TGTCTGTAGCTCAACTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2400		; PRIOR APPLICATION NUMBER: US 09/332,782			
Db	4360	TGTCTGTAGCTCAACTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	4419		; PRIOR FILING DATE: 1999-06-14			
Qy	2401	TCCACATCTTGCACAGACAGGCTTTTGTCCAGCTGTCACATTTGAGTGCAGCTGCC	2460		; PRIOR APPLICATION NUMBER: US 09/235,076			
Db	4420	TCCACATCTTGCACAGACAGGCTTTTGTCCAGCTGTCACATTTGAGTGCAGCTGCC	4478		; PRIOR FILING DATE: 1999-01-20			
Qy	2461	GGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2520		; PRIOR APPLICATION NUMBER: US 09/234,611			

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122)...(2969)
; OTHER INFORMATION: similar to gil841434 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTX 3.3t00, default parameters
US-09-496-914A-6881

Query Match 68.1%; Score 1790.4; DB 22; Length 4487;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 2297; Conservative 5; Mismatches 190; Indels 87; Gaps 34;

Qy	1	CCCTCTCACAGCCAGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCACCGCTGT	60
Db	1993	CCCTCTCACAGCCAGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCACCGCTGT	2052
Qy	61	CTGAGCGGAGCTGGGAGGAGGTGAACCGGCACTACAGCAAGTGGAGGTCTGAAGA	120
Db	2053	CTGAGCGGAGCTGGGAGGAGGTGAACCGGCACTACAGCAAGTGGAGGTCTGAAGA	2112
Qy	121	GCCCTTGAGGGGAGATATAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATACC	180
Db	2113	GCCCTTGAGGGGAGATATAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATACC	2172
Qy	181	ACCAGACCTTCATGCCAGCCGAGAGAGCTTTCCCAAGGGCCCGAGGGCCCGGCCAG	240
Db	2173	ACCAGACCTTCATGCCAGCCGAGAGAGCTTTCCCAAGGGCCCGAGGGCCCGGCCAG	2232
Qy	241	CTGAGGAGACAAAGGAGAGCCCTTAAGTCTCAGCTCTCTCCACACGAGCCCGCCAG	300
Db	2233	CTGAGGAGACAAAGGAGAGCCCTTAAGTCTCAGCTCTCTCCACACGAGCCCGCCAG	2292
Qy	301	AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAGGAGGAGTCTGGAGTGGGAAACCT	360
Db	2293	AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAGGAGGAGTCTGGAGTGGGAAACCT	2352
Qy	361	TACCTCTGTCTCTCTGGAGCAGCCCTCTGCAGAAACCCAGCTCACAGAGCGGAAAG	420
Db	2353	TACCTCTGTCTCTCTGGAGCAGCCCTCTGCAGAAACCCAGCTCACAGAGCGGAAAG	2412
Qy	421	CAACCGTCCGGAGCAGGAATGAGCAGCTGGAATAGAAATTTCTCTCAACAGCCTGT	480
Db	2413	CAACCGTCCGGAGCAGGAATGAGCAGCTGGAATAGAAATTTCTCTCAACAGCCTGT	2472
Qy	481	CCAGCCATTTCTCTGAGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCC	540
Db	2473	CCAGCCATTTCTCTGAGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCC	2532
Qy	541	TCTCCCTGTCCGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGGCTCGCGGACA	600
Db	2533	TCTCCCTGTCCGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGGCTCGCGGACA	2592
Qy	601	CCCTGAGCTCAGGCGTACCTCTGGAGAGCCAGGCGGAGGTCTGAAGTCTCAGCTGGA	660
Db	2593	CCCTGAGCTCAGGCGTACCTCTGGAGAGCCAGGCGGAGGTCTGAAGTCTCAGCTGGA	2652
Qy	661	ACATGGTCTGGCCCGGGGGGGCCACGACACCCCAAGCTATTCAATGGTGAAG	720
Db	2653	ACATGGTCTGGCCCGGGGGGGCCACGACACCCCAAGCTATTCAATGGTGAAG	2712
Qy	721	TCCAAATACAGTCTCTTAATGGTGAACAACCTGACATCCGGGAGTTCACCGGTCAAG	780
Db	2713	TCCAAATACAGTCTCTTAATGGTGAACAACCTGACATCCGGGAGTTCACCGGTCAAG	2772
Qy	781	TGGAGACATCGGCATCAGCAGCCAGATCCAGCTGACGCTTCAAGCTTGGTCA	840
Db	2773	TGGAGACATCGGCATCAGCAGCCAGATCCAGCTGACGCTTCAAGCTTGGTCA	2832
Qy	841	CCAAAGACGGGACCTTGTCTGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC	900
Db	2833	CCAAAGACGGGACCTTGTCTGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC	2892

Qy	901	AGTGCACACTGGCCCTTGATGGCAGCTTCGCTGGAGCTGGAGGGTCAAGCATGGCCAGC	960
Db	2893	AGTGCACACTGGCCCTTGATGGCAGCTTCGCTGGAGCTGGAGGGTCAAGCATGGCCAGC	2952
Qy	961	TGAGAAACAGCCCTAACCTGCGCTCCACCGCGCTCCACACTGCGGAAAGACAGCCT	1020
Db	2953	TGAGAAACAGCCCTAACCTGCGCTCCACCGCGCTCCACACTGCGG--AAGACAGCCT	3011
Qy	1021	TCCTGCTCGGTGACCATGCTGCGCTGAAAACACAGGCTCAGCCGTTCCAGGGGATVTG	1080
Db	3012	TCCTGCTCGGTGACCATGCTGCGCTG--AAACACAGGCTCAGCCGTTCCAGGG--ATC	3067
Qy	1081	NCAGACCCCGCGCTCARCAGNTGGGAAACACAGGCGCTTCGNCAGNAGNAGTNGGGG	1140
Db	3068	TGCAGCCCGCGCTCAGCAGTGGGACACAGGCGCTTCGACAGCAGCAAG-----GTGGG	3120
Qy	1141	CAAGCAGATGCTCCCGAGATTTCACANCCTGAGCCNTGCCANCCCTGCTGAGADA	1200
Db	3121	GCAAGCAGATGCTCCCGAGATTTCACAC--TGAGCCCTGCCCCACCCCTGCTGAGAA	3177
Qy	1201	AAACAYTNCCGCCACGTGAAGAGACAGAGGAGATGNCAGGAGTTNNACCTTGGGGAA	1260
Db	3178	AAACT--CCGCCACGTGAAGAGACA---GAGGAGATGGCAGGAGTTACCTCGGANA	3231
Qy	1261	ACAAAACAGGGATCTTTNTTCTGCCCTGCTCCAGTNCAGTGGCTGCGNACCCTTGG	1320
Db	3232	CAAAACAGGATCTTCTGCTGCTGCTCAG-----TCGAGTTGGCTGACCCGCTT--G	3282
Qy	1321	ANTCAGTGACCATTTGTTGGCAGANCAGGGGAGAGAGCTTCCAGCTTGGTCAAGGG	1380
Db	3283	GATCAGTGACCATTTGTTGGCAGA-CAGGGGAGAGAGCTTCCAGCTTGGTCAAGAGGG	3341
Qy	1381	GTGGGCGAGCCTTCGGCCCTCAACCTNCAGGCTCTGTGNAGAGTGTCAAGTGTGTA	1440
Db	3342	GTGGGCGAGCCTTCGGCCCTCAACCT--CCAGCTGCTGTG--AGAGTGTCAAGTGTGTA	3399
Qy	1441	AGGNNCCAAANTCAGGNTTCAGTGCAGAACACAGGTCNACAGAGGTATGCCGCCCGNTA	1500
Db	3400	AGGGCCCAAACTC---AGGTTTCAGTGCAGAACACAGGT-CAGCAGGTATGCCGCCCGTAG	3455
Qy	1501	GGTTAANNNGGGGCGCTCTNAAACCCCTTGCTNGGCTNCACCTNGGCCAGCTCANCCC	1560
Db	3456	GTATA---GGGGGCGCTCTNAAACCCCTTGCTGCGCTCA---CCTGGCCAGCTCA--CCC	3507
Qy	1561	CTTTTGGGTGTAGGGGAAAGAAATGCTGACCTCGGGAAGGCTTWCCTCTGGTAGAATACAC	1620
Db	3508	CTTTTGGGTGTAGGGGAAAGAAATGCTGACCTCGGGAAGGCT--CCCTGGTAGAATACAC	3566
Qy	1621	CACACTTTTTCAGGTTGTGCAACACAGGCTCTGAGTTGACCTCTGGTTCAGCCAGGACC	1680
Db	3567	CACACTTTTTCAGGTTGTGCAACACAGGCTCTGAGTTGACCTCTGGTTCAGCCAGGACC	3626
Qy	1681	AAAGAGGTGTGTAAGTGAAGTGTCTCAGTNCCTCAGACATGTGCCCTTTGCTGCTGTG	1740
Db	3627	AAAGAGGTGTGTAAGTGAAGTGTCTCAGT--CCCCAGACATGTGCCCTTTGCTGCTGTG	3685
Qy	1741	GCTTACCACCTTCCCGCAGACAGCGCCCGGAGCCCTTTCAGGCCAGACACTGCCCCAG	1800
Db	3686	GCTTACCACCTTCCCGCAGACAGCGCCCGGAGCCCTTTCAGGCCAGACACTGCCCCAG	3745
Qy	1801	ACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGAGGATATGCTGTA	1860
Db	3746	ACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGAGGATATGCTGTA	3805
Qy	1861	CAGGAAACAGTCTGTGATGGACATGATCAGTGTNAAAGNAAAGACAGACAGAGAGAGY	1920
Db	3806	CAGGAAACAGTCTGTGATGGACATGATCAGTGTGTA--GGAAAACAGACAGAGAGAGC	3863
Qy	1921	TCCGGCGCCCGCCAGNCCCACTNATCAGTGTNCCAGGCTGCTNGTTCNCCCGAGNAGACA	1980
Db	3864	TCCGGCGCCCGCCAGNCCCACTNATCAGTGT---CCAGGCTGCTGCTTCC---CAGA	3912
Qy	1981	GCTNAGNATCANCACTGACACTNCACTNCCCTGCTGCCCTNGGCCANGAGGGTACTG	2040

Mon Jun 13 13:27:12 2005

Db	3913	GCACAGCTCAGCATCACACTGACCTACCCCTGCCCTGCCCTGG--CCAGAGGGTACTG	3970
Qy	2041	CCGNACGGCACTTTGGCACTCTGATGATGACCTCAAGCACTTTCATGGCTNGCCCTCTNNG	2100
Db	3971	CCG-ACGGCACTTTGCACT--CTGATGACCTCAAGCACTTTCATGGCTGCCCTCTGGCA	4027
Qy	2101	GCAGGNCAGGNCAGGNCAGTGCACNCTGTAGNAGCATANGCAANGCCAGGAGATGG	2160
Db	4028	GGGACAGGAGGGCAGTACACT-----GTAGGAGCATAGCAAGCCAGGAGATGG	4077
Qy	2161	GGTNAAGGGANACAGTCTTGAGTGTCCANCATGATGTGACTNCTCAAACTCTTN	2220
Db	4078	GGTGA--GGGACACAGTCTTGAGTGTCCA-CATGCAATGTGACT-CCTCAAACTCTTT-	4132
Qy	2221	NCCAGNATTTCTTAAGATAGCANCCTTNCCTTCCCTGCTTCCAGCTTAGCCCTCTTCT	2280
Db	4133	--CCAGATTTCTTAAGATAGCACCCTC--TTCCCTTGGCCAGCTTAGCCCTCTTCT	4188
Qy	2281	CCAGGGGAGCTANCTCAGGACTCACGTAGCATTAATCAGCTGTGNAATCGTCAGGGGG	2340
Db	4189	CCAGGGGAGCTA-CTCAGGACTCACGTAGCATTAATCAGCTGTG-AATCGTCAGGGGG	4246
Qy	2341	TGTCCTAGCCTCAACCTCTGGGGCAGGGACGCGAGACTCCGTTGGGAGAGCTCAT	2400
Db	4247	TGTCCTAGCCTCAACCTCTGGGGCAGGGACGCGAGACTCCGTTGGGAGAGCTCAT	4306
Qy	2401	TCCACATCTTGCCAAAGACAGCCTTTNGTCCAGCTGTCCACATTCAGTCAGACTGCTCCC	2460
Db	4307	TCCACATCTTGCCAAAGACAGCCTTT-GTCCAGCTGTCCACATTCAGTCAGACTGCTCCC	4365
Qy	2461	GGGAGAGAGCCCCGGCCCCAGCACATAAAGAACTGAGCCCTTGGTACTGCGAGTCTG	2520
Db	4366	GGGAGAGAGCCCCGGCCCCAGCACATAAAGAACTGAGCCCTTGGTACTGCGAGTCTG	4425
Qy	2521	GTTGTAGAACTCTTTGTAAGCAATAAAGTTTGGGGTGTGACAAAAATGTTAAAAAAA	2579
Db	4426	GTTGTAGAACTCTTTGTAAGCAATAAAGTTTGGGGTGTGACAAAAATGTTAAAAAAA	4484

Search completed: June 11, 2005, 04:00:58
Job time : 6603.79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 09:44:05 ; Search time 1006.45 Seconds
(without alignments)
11471.453 Million cell updates/sec

Title: US-09-155-676B-3
Perfect score: 2631
Sequence: 1 cccctctcacgccagcc.....accgtgacgtcgaggsgggg 2631

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 12985247 seqs, 2194111492 residues

Total number of hits satisfying chosen parameters: 25970494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA New:
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	2553.8	97.1	2631	6	US-09-155-676B-3	Sequence 3, Appli
2	1851.2	70.4	4596	6	US-09-155-676B-6	Sequence 6, Appli
3	1850.2	70.3	4596	8	US-10-450-763-15294	Sequence 15294, A
4	1781.4	67.7	4458	15	US-60-659-397-567	Sequence 567, App
5	1781.4	67.7	4684	15	US-60-659-397-566	Sequence 566, App
6	971.8	36.9	65967	15	US-60-659-397-12129	Sequence 12129, A
7	769.2	29.2	1400	13	US-11-060-756-560	Sequence 560, App
8	769.2	29.2	1400	13	US-11-060-756-4832	Sequence 4832, Ap
9	691.8	26.3	2760	8	US-10-450-763-15295	Sequence 15295, A
10	588.2	22.4	722	8	US-10-450-763-15296	Sequence 15296, A
11	295.6	11.2	3658	15	US-60-680-544-32424	Sequence 32424, A
12	295.6	11.2	3658	15	US-60-680-473-32424	Sequence 32424, A
13	257.4	9.8	761	15	US-60-680-544-10785	Sequence 10785, A
14	257.4	9.8	761	15	US-60-680-544-11056	Sequence 11056, A
15	257.4	9.8	761	15	US-60-680-473-10785	Sequence 10785, A
16	257.4	9.8	761	15	US-60-680-473-11056	Sequence 11056, A
17	200.6	7.6	201	15	US-60-659-397-7882	Sequence 7882, Ap
18	200.6	7.6	201	15	US-60-659-397-7887	Sequence 7887, Ap

19 200.6 7.6 201 15 US-60-659-397-7888 Sequence 7888, Ap
20 200.6 7.6 201 15 US-60-659-397-7893 Sequence 7893, Ap
21 189.6 7.2 201 15 US-60-659-397-7885 Sequence 7885, Ap
22 189.6 7.2 201 15 US-60-659-397-7891 Sequence 7891, Ap
23 189.6 7.2 201 15 US-60-659-397-32724 Sequence 32724, A
24 188.6 6.4 201 15 US-60-659-397-7886 Sequence 7886, Ap
25 188.6 6.4 201 15 US-60-659-397-7892 Sequence 7892, Ap
26 168.6 6.4 201 15 US-60-659-397-32758 Sequence 32758, A
27 150 5.7 201 15 US-60-659-397-32751 Sequence 32751, A
28 128.6 4.9 201 15 US-60-659-397-32755 Sequence 32755, A
29 119.2 4.5 1353 15 US-60-680-544-37753 Sequence 37753, A
30 119.2 4.5 1353 15 US-60-680-473-37753 Sequence 37753, A
31 96.4 3.7 201 15 US-60-659-397-7884 Sequence 7884, Ap
32 96.4 3.7 201 15 US-60-659-397-7890 Sequence 7890, Ap
33 96.4 3.7 201 15 US-60-659-397-7857 Sequence 7857, A
34 88 3.3 201 15 US-60-659-397-32726 Sequence 32726, A
35 78 3.0 201 15 US-60-659-397-32752 Sequence 32752, A
36 78 3.0 599 9 US-10-972-079-7876 Sequence 7876, Ap
37 70 2.7 1024 8 US-10-450-763-15297 Sequence 15297, A
38 68.8 2.6 201 15 US-60-659-397-32718 Sequence 32718, A
39 62.4 2.4 600 9 US-10-972-079-7877 Sequence 7877, Ap
40 59.4 2.3 493 1 PCT-US05-00517-813 Sequence 813, App
41 53.2 2.0 1253 6 US-09-155-676B-4 Sequence 4, Appli
42 53.2 2.0 1253 13 US-11-060-756-1553 Sequence 1553, Ap
43 53.2 2.0 1253 13 US-11-060-756-1554 Sequence 1554, Ap
44 53.2 2.0 1253 13 US-11-060-756-5825 Sequence 5825, Ap
45 53.2 2.0 1253 13 US-11-060-756-5826 Sequence 5826, Ap

ALIGNMENTS

RESULT 1

US-09-155-676B-3
; Sequence 3, Application US/09155676B
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR
; FILE OF INVENTION: PREPARATION AND USE
; FILE REFERENCE: WALLACH-21
; CURRENT APPLICATION NUMBER: US/09/155.676B
; CURRENT FILING DATE: 1999-01-04
; PRIOR APPLICATION NUMBER: PCT/IL97/00117
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: IL 117800
; PRIOR FILING DATE: 1996-04-02
; PRIOR APPLICATION NUMBER: IL 119133
; PRIOR FILING DATE: 1996-08-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 3
; LENGTH: 2631
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1081)..(1081)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1102)..(1102)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1120)..(1120)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1125)..(1125)

Mon Jun 13 13:27:13 2005

us-09-155-676b-3.rnpn

OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1129)..(1129)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1135)..(1135)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1146)..(1146)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1170)..(1170)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1180)..(1180)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1188)..(1188)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1208)..(1208)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1239)..(1239)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1248)..(1249)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1278)..(1278)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1297)..(1297)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1310)..(1310)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1322)..(1322)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1345)..(1345)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1409)..(1409)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1423)..(1423)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1445)..(1445)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1452)..(1452)
OTHER INFORMATION: n is a, c, g, or t

FEATURE:
NAME/KEY: misc feature
LOCATION: (1459)..(1459)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1478)..(1478)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1498)..(1498)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1507)..(1508)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1520)..(1520)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1534)..(1534)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1540)..(1540)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1546)..(1546)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1557)..(1557)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1713)..(1713)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1895)..(1895)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1900)..(1900)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1934)..(1934)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1942)..(1942)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1951)..(1951)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1962)..(1962)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1967)..(1967)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (1974)..(1974)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:

```
; NAME/KEY: misc feature
; LOCATION: (1984)..(1984)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1988)..(1988)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1994)..(1994)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2005)..(2005)
; OTHER INFORMATION: n is a, c, g, or t

Query Match 97.1%; Score 2553.8; DB 6; Length 2631;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCCTCTCACAGCCAGGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
Db 1 CCCCTCTCACAGCCAGGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60

Qy 61 CTGAGGAGGAGCTGGAGGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120
Db 61 CTGAGGAGGAGCTGGAGGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120

Qy 121 GCCCTTGGAGGGAGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 180
Db 121 GCCCTTGGAGGGAGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 180

Qy 181 ACCAGACCTCTCATGCCAGCCGAGAGAGCTTTTCGCCAAAGGGCCCGAGGGCCCGGCCAG 240
Db 181 ACCAGACCTCTCATGCCAGCCGAGAGAGCTTTTCGCCAAAGGGCCCGAGGGCCCGGCCAG 240

Qy 241 CTGAGGAGACAAGGAGAGAGCCCTTAAGTCTCAGCTCTCTCCACCAAGAGCCCGCCAG 300
Db 241 CTGAGGAGACAAGGAGAGAGCCCTTAAGTCTCAGCTCTCTCCACCAAGAGCCCGCCAG 300

Qy 301 AGCCAAACAGTCTCTCCCTTGACTTTTGACAGAGGAGGAGTCTGGATCTGGGAACCTT 360
Db 301 AGCCAAACAGTCTCTCCCTTGACTTTTGACAGAGGAGGAGTCTGGATCTGGGAACCTT 360

Qy 361 TACCTCTGTCTCTCTGGAGCAGCCCTTGCCAGAAACCCAGCTCACAGAGCGGAAAG 420
Db 361 TACCTCTGTCTCTCTGGAGCAGCCCTTGCCAGAAACCCAGCTCACAGAGCGGAAAG 420

Qy 421 CAACCGTCCCGAGCAGGAACTGAGCAGCTGGAAATAGAAATATTCTCAACAGCCTGT 480
Db 421 CAACCGTCCCGAGCAGGAACTGAGCAGCTGGAAATAGAAATATTCTCAACAGCCTGT 480

Qy 481 CCCAGCATTTTCTCTGGAGAGCAGAGCAAAATCTCTGTGCTCAGCATCAGCAGCC 540
Db 481 CCCAGCATTTTCTCTGGAGAGCAGAGCAAAATCTCTGTGCTCAGCATCAGCAGCC 540

Qy 541 TCTCCCTGTGCGATGACAGTGAGAAAGAACCCATCAAAAGGCTCTCAAAAGCTCGGGGACA 600
Db 541 TCTCCCTGTGCGATGACAGTGAGAAAGAACCCATCAAAAGGCTCTCAAAAGCTCGGGGACA 600

Qy 601 CCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTTCCAGCTGGA 660
Db 601 CCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTTCCAGCTGGA 660

Qy 661 ACATGTGTGCGCGGGGGGGCCCAACCGACACCCCAAGCTATTTCATGTGTGTGAAG 720
Db 661 ACATGTGTGCGCGGGGGGGCCCAACCGACACCCCAAGCTATTTCATGTGTGTGAAG 720

Qy 721 TCCAAATACAGTCTCTTAATGGTGAACACCTTGCACATCCGGGAGTTCCACCGGTCAAAG 780
Db 721 TCCAAATACAGTCTCTTAATGGTGAACACCTTGCACATCCGGGAGTTCCACCGGTCAAAG 780

Qy 781 TGGAGAGACATCGGCATCAGCAGCCAGATCCCGAGCTTCCAGCTTGTGTCA 840
Db 781 TGGAGAGACATCGGCATCAGCAGCCAGATCCCGAGCTTCCAGCTTGTGTCA 840

841 CCAAAGACGGGAGCCTGTTTCGCTACGACATGAGAGTGCCAGACTCGGGCATCGACCTGC 900
841 CCAAAGACGGGAGCCTGTTTCGCTACGACATGAGAGTGCCAGACTCGGGCATCGACCTGC 900

901 AGTGCACTGGGCGCCCTGATGGGAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 960
901 AGTGCACTGGGCGCCCTGATGGGAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 960

961 TGGAGACAGCCCTTAACCTGCGCTCCACGGCGGCTCCACATGCGCGGAAGCAGCCT 1020
961 TGGAGACAGCCCTTAACCTGCGCTCCACGGCGGCTCCACATGCGCGGAAGCAGCCT 1020

1021 TCCTGCTCGGTGACAGATGCTGCCCTGAAAACACAGGCTCAGCCGCTTCCAGGGGATYTG 1080
1021 TCCTGCTCGGTGACAGATGCTGCCCTGAAAACACAGGCTCAGCCGCTTCCAGGGGATYTG 1080

1081 NCCAGCCCCCGGCTCARCAGNTGGGAAACAGGGCCTCGNCAGCNAGCNAAAGTNGGGG 1140
1081 NCCAGCCCCCGGCTCARCAGNTGGGAAACAGGGCCTCGNCAGCNAGCNAAAGTNGGGG 1140

1141 CAAGCNAGATGCTCCAGGATTTTTCANCTGAGCCCTGCCANCCCTGCTGAADA 1200
1141 CAAGCNAGATGCTCCAGGATTTTTCANCTGAGCCCTGCCANCCCTGCTGAADA 1200

1201 AAACAYTNCGCCAGCTGAAGAGAGAGAGATGNCAGAGGATGNCAGGATGNNACCTYGGGAA 1260
1201 AAACAYTNCGCCAGCTGAAGAGAGAGAGATGNCAGAGGATGNCAGGATGNNACCTYGGGAA 1260

1261 ACAAACAGGAGATTTTNTTTCCTGCTTCCAGTNCAGGATTTGGCCTTGNACCCGCTTGG 1320
1261 ACAAACAGGAGATTTTNTTTCCTGCTTCCAGTNCAGGATTTGGCCTTGNACCCGCTTGG 1320

1321 ANTCACTGACATTTTGTGGCAGANCAGGGGAGAGAGCTTCCAGCTTGGGTCAAGAGGG 1380
1321 ANTCACTGACATTTTGTGGCAGANCAGGGGAGAGAGCTTCCAGCTTGGGTCAAGAGGG 1380

1381 GTGGGCGAGCCCTTTCGGCCCTTCCAGCTTCCAGGCTGCTGTGNAGAGTGTCAAGTGTGA 1440
1381 GTGGGCGAGCCCTTTCGGCCCTTCCAGCTTCCAGGCTGCTGTGNAGAGTGTCAAGTGTGA 1440

1441 AGGNCNCCAAANCTCAGGNTTCAGTGAGAAACAGAGTNCAGCAGGATATGCCCCCGGNTA 1500
1441 AGGNCNCCAAANCTCAGGNTTCAGTGAGAAACAGAGTNCAGCAGGATATGCCCCCGGNTA 1500

1501 GGTAAANNNGGGGGCCCTTNAACCCCTTGGCTNGGCTTNCACCTNGGCGCAGCTCANCCC 1560
1501 GGTAAANNNGGGGGCCCTTNAACCCCTTGGCTNGGCTTNCACCTNGGCGCAGCTCANCCC 1560

1561 CTTTGGGTGTAGGGGAAAAGATGCTGACCTTGGGAAGGCTWCCCTGGTAGAATACAC 1620
1561 CTTTGGGTGTAGGGGAAAAGATGCTGACCTTGGGAAGGCTWCCCTGGTAGAATACAC 1620

1621 CACACTTTTTCAGGTTGTGCAACACAGGCTCTGAGTTGACCTCTGTTTCAGCCHAGGACC 1680
1621 CACACTTTTTCAGGTTGTGCAACACAGGCTCTGAGTTGACCTCTGTTTCAGCCHAGGACC 1680

1681 AAAGAAGGTGTGAAGTGAAGTGTCTCATGATNCCCAGACATGTGCCCCCTTGTGCTGTG 1740
1681 AAAGAAGGTGTGAAGTGAAGTGTCTCATGATNCCCAGACATGTGCCCCCTTGTGCTGTG 1740

1741 GCTTACCACTTCTCCAGAGCAGGCGCCCGAGCCCTTTCAGGCCAGCACTGCCCCAG 1800
1741 GCTTACCACTTCTCCAGAGCAGGCGCCCGAGCCCTTTCAGGCCAGCACTGCCCCAG 1800

1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAGAGTGAAGGTGATGAGGATATGCTCTGA 1860
1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAGAGTGAAGGTGATGAGGATATGCTCTGA 1860

1861 CAGGAACAGTCTGTGGATGGAATGATGATGCTGCTTAAAGGNAAGCAGCAGAGAGAGGY 1920
1861 CAGGAACAGTCTGTGGATGGAATGATGATGCTGCTTAAAGGNAAGCAGCAGAGAGAGGY 1920
```


QY	61	CTCAGCGGAGCTGGAGGGGAAGGTGAA	CGGGGCACTACAGCAAGTGGAGGTCTGAAGA	120
Db	2160	CTCAGCGGAGCTGGAGGGGAAGGTGAA	CGGGGCACTACAGCAAGTGGAGGTCTGAAGA	2219
QY	121	GCCTTTGGAGGGGAGAAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC	180	
Db	2220	GCCTTTGGAGGGGAGAAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC	2279	
QY	181	ACAGACCCCTCCATGCCCGACCGAGAGAGGTTTCGGCCAAAGGGCCCGACGGCCCGGCCAG	240	
Db	2280	ACAGACCCCTCCATGCCCGACCGAGAGAGTTTCGGCCAAAGGGCCCGACGGCCCGGCCAG	2339	
QY	241	CTGAGGAGACAACAGGAGAGCCCTTAAGTCTCAGGCTCTCTCCACAGAGCCCGCCAG	300	
Db	2340	CTGAGGAGACAACAGGAGAGCCCTTAAGTCTCAGGCTCTCTCCACAGAGCCCGCCAG	2399	
QY	301	AGCCAAACAAAGTCTCCCTTCGACTTTGAGCAAGGAGGAGTCTGGGAATGTGGGAACCCCT	360	
Db	2400	AGCCAAACAAAGTCTCCCTTCGACTTTGAGCAAGGAGGAGTCTGGGAATGTGGGAACCCCT	2459	
QY	361	TACCTCTGTCTCCTCGAGGAGCAGCCCTCCAGAAACCCAGCTCACAGAGCGGAAG	420	
Db	2460	TACCTCTGTCTCCTCGAGGAGCAGCCCTCCAGAAACCCAGCTCACAGAGCGGAAG	2519	
QY	421	CAACCGTCCCGAGCAGGAATGACAGCTGAGGAGTGGAAATAGAAATATTCTCAACAGAGCTGT	480	
Db	2520	CAACCGTCCCGAGCAGGAATGACAGCTGAGGAGTGGAAATAGAAATATTCTCAACAGAGCTGT	2579	
QY	481	CCAGGCCAATTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGCCTCAGCATCGACAGCC	540	
Db	2580	CCAGGCCAATTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGCCTCAGCATCGACAGCC	2639	
QY	541	TCTCCTGTTCGATGACAGTGAAGAAACCCATCAAAAGGCTCTCAAAAGCTCCGGGACA	600	
Db	2640	TCTCCTGTTCGATGACAGTGAAGAAACCCATCAAAAGGCTCTCAAAAGCTCCGGGACA	2699	
QY	601	CCCTGAGCTCAGGCTGACACTCTCGAGCAGCCGAGGCTCGAAGCTCCAGTGTGA	660	
Db	2700	CCCTGAGCTCAGGCTGACACTCTCGAGCAGCCGAGGCTCGAAGCTCCAGTGTGA	2759	
QY	661	ACATGGTGTGCGCCGGGGGGGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG	720	
Db	2760	ACATGGTGTGCGCCGGGGGGGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG	2819	
QY	721	TCCAAATACAGTCTCTTAATGTGTGAACCTCGACATCCCGGGTCCACCGGCTCAAAG	780	
Db	2820	TCCAAATACAGTCTCTTAATGTGTGAACCTCGACATCCCGGGTCCACCGGCTCAAAG	2879	
QY	781	TGGGAGACATCGCCATCGGCATCAGCAGCCAGATCCACAGCTGCAGCTTCAGCTTGTGCA	840	
Db	2880	TGGGAGACATCGCCATCGGCATCAGCAGCCAGATCCACAGCTGCAGCTTCAGCTTGTGCA	2939	
QY	841	CCAAAGACGGGAGCCTGTTTCGCTACGACATGGAGGTTGCAGATCCGGGCAATCGACCTGC	900	
Db	2940	CCAAAGACGGGAGCCTGTTTCGCTACGACATGGAGGTTGCAGATCCGGGCAATCGACCTGC	2999	
QY	901	AGTGTGACATCGGCCCCGTGATGGAGCTTCGCTGAGAGTGGAGGTTCAAGCATGGCCAGC	960	
Db	3000	AGTGTGACATCGGCCCCGTGATGGAGCTTCGCTGAGAGTGGAGGTTCAAGCATGGCCAGC	3059	
QY	961	TGGAGAACAGGCCCCCTAACCTCGCCCTCCACCGCGGCTCCACATGCGCGGAAGCAGCCT	1020	
Db	3060	TGGAGAACAGGCCCCCTAACCTCGCCCTCCACCGCGGCTCCACATGCGCGGAAGCAGCCT	3119	
QY	1021	TCCTGTCTCGTGCAGATGTCCTGTGAATAACAGGCTCAGCGCTCCCGGGGATGTG	1080	
Db	3120	TCCTGTCTCGTGCAGATGTCCTGTGAATAACAGGCTCAGCGCTCCCGGGGATGTG	3177	
QY	1081	NCCAGCCCCCGGCTCARGCAGNTGGGAACCCAGGCGCTCGNCAGCAGCAGGTTGGGGG	1140	
Db	3178	GCCAGCCCCCGGCTCARGCAGNTGGGAACCCAGGCGCTCGNCAGCAGCAGGTTGGGGG	3230	
QY	1141	CAAGCNAGAAATGCTCCAGGAATTCACANCCCTGAGCCNTGCCCCNCCCTGCTGAADA	1200	

Db 4246 --CCAGATTCTTAAGAAATAGACCCCC--TTCCCCATTCGCCAGCTTAGCCTCTTCT 4301
Qy 2281 CCCAGGGAGTANCTCAGGACTCAGCTAGCAATTAATCAGCTGTGNAATCGTCAGGGGG 2340
Db 4302 CCCAGGGAGCTA-CTCAGGACTCAGCTAGCAATTAATCAGCTGTG-AATCGTCAGGGGG 4359
Qy 2341 TGTCTGTAGCTCAACCTCTCTGGGCGAGGGAGCGCCGAGACTCCGTGGGAGAAAGCTCAT 2400
Db 4360 TGTCTGTAGCTCAACCTCTCTGGGCGAGGGAGCGCCGAGACTCCGTGGGAGAAAGCTCAT 4419
Qy 2401 TCCACATCTTTGCCAAGACAGCTTTTNGTCCAGCTGTCCACATTAAGTTCAGACTGCTCCC 2460
Db 4420 TCCACATCTTTGCCAAGACAGCTTTT-GTCCAGCTGTCCACATTAAGTTCAGACTGCTCCC 4478
Qy 2461 GGGGAGAGAGCCCGGGCCCCCAGCACATAAAGAACTGCAGCTTTGGTACTGCAGAGCTG 2520
Db 4479 GGGGAGAGAGCCCGGGCCCCCAGCACATAAAGAACTGCAGCTTTGGTACTGCAGAGCTG 4538
Qy 2521 GGTGTAGAGAACTCTTTTGAAGCAATAAAGTTTGGGGTGAAGCAAAATGTTAAAAA 2578
Db 4539 GGTGTAGAGAACTCTTTTGAAGCAATAAAGTTTGGGGTGAAGCAAAATGTTAAAAA 4596

RESULT 4
US-60-659-397-567
; Sequence 567, Application US/60659397
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/659,397
; CURRENT FILING DATE: 2005-03-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 567
; LENGTH: 4458
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-659-397-567

Query Match 67.7%; Score 1781.4; DB 15; Length 4458;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 2292; Conservative 10; Mismatches 183; Indels 87; Gaps 35;

Qy 1 CCCCTCTCACAGCCAGCCATCCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
Db 1974 CCCCTCTCACAGCCAGCCATCCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 2033
Qy 61 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA 120
Db 2034 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA 2093
Qy 121 GCCCTTGGAGGGGAGATATATAAGAACCAAGACATCCACGCCCAATCAAGCCAAATTACC 180
Db 2094 GCCCTTGGAGGGGAGATATATAAGAACCAAGAYATCCACGCCCAATCAAGCCAAATTACC 2153
Qy 181 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCGGCCAG 240
Db 2154 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCGGCCAG 2213
Qy 241 CTGAGGAGACAAACAGGAGAGCCCTTAAGTCTCAGCTCTCTCCACCAAGAGCCCCCAG 300
Db 2214 CTGAGGAGACAAACAGGAGAGCCCTTAAGTCTCAGCTCTCTCCACCAAGAGCCCCCAG 2273
Qy 301 AGCCAAACAGTCTCTCCCTTGAATTTGACCAAGAGGAGTCTGGGATGGAACCCCT 360
Db 2274 AGCCAAACAGTCTCTCCCTTGAATTTGACCAAGAGGAGTCTGGGATGGAACCCCT 2333
Qy 361 TACCTCTGCTCTCCCTGGAGGCCCCCTTGCAGAAACCCAGCTCACCAGAGCGGAAG 420

Db 2334 TACCTCTGCTCTCCCTGGAGCCAGCCCTCCAGAAACCCAGCTCACCAGAGCGGAAG 2393
Qy 421 CAACCGTCCCGGAGCAGGAACCTGCAGCAGCTGGAATAGAAATATTCTCTCAACAGCCTGT 480
Db 2394 CAACCGTCCCGGAGCAGGAACCTGCAGCAGCTGGAATAGAAATATTCTCTCAACAGCCTGT 2453
Qy 481 CCAGAGCAATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCC 540
Db 2454 CCAGAGCAATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCC 2513
Qy 541 TCTCCCTGTCCGATGACAGTGAAGAACCCATCAAAAGGCTCTCAAAAGCTCCGCGGACA 600
Db 2514 TCTCCCTGTCCGATGACAGTGAAGAACCCATCAAAAGGCTCTCAAAAGCTCCGCGGACA 2573
Qy 601 CCCTGAGCTCAGCGGTACACTCTCTGGAGCAGCAGGCGGAGGCTCGAAGCTCCAGCTTGA 660
Db 2574 CCCTGAGCTCAGCGGTACACTCTCTGGAGCAGCAGGCGGAGGCTCGAAGCTCCAGCTTGA 2633
Qy 661 ACATGCTGTGGCCCGGGGCGGCCCAACGACACCCCAAGCTATTTCATATGCTGTAAG 720
Db 2634 ACATGCTGTGGCCCGGGGCGGCCCAACGACACCCCAAGCTATTTCATATGCTGTAAG 2693
Qy 721 TCCAAATACAGTCTCTTAATGTTGAACACCTGCACATCCGGGAGTTCCACCGGTCAAAG 780
Db 2694 TCCAAATACAGTCTCTTAATGTTGAACACCTGCACATCCGGGAGTTCCACCGGTCAAAG 2753
Qy 781 TGGGAGACATCGCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTCAGCTTGTGTC 840
Db 2754 TGGGAGACATCGCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTCAGCTTGTGTC 2813
Qy 841 CCAAAGACGGGCGGCGCTTTCGCTACGACATGGAGGTGCAGACTCCGGGATCGACCTGC 900
Db 2814 CCAAAGACGGGCGGCGCTTTCGCTACGACATGGAGGTGCAGACTCCGGGATCGACCTGC 2873
Qy 901 AGTGACACTGGCCCTGTATGGCAGCTTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGC 960
Db 2874 AGTGACACTGGCCCTGTATGGCAGCTTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGC 2933
Qy 961 TGGAGAACAGGCGCTTAACCTGCGCTCCACGCGCGCTCCACACTGCCGGAAGACAGCCT 1020
Db 2934 TGGAGAACAGGCGCTTAACCTGCGCTCCACGCGCGCTCCACACTGCCGGAAGACAGCCT 2992
Qy 1021 TCCTGCTCGGTGACATGCTGCCCTGAAACACAGGCTCAGCGCTTCCAGGGGATGTC 1080
Db 2993 TCCTGCTCGGTGACATGCTGCCCTG- AAACACAGGCTCAGCGCTTCMAGGG- --ATC 3048
Qy 1081 NCCAGCCCCCGGCTCARCAGNTGGGAACACAGGGCTTCGNCAGCAGCNAAGTNGGGGG 1140
Db 3049 TGCCAGCCCCCGGCTCAGCAGTGGGACCCAGGGCTCGCAGCAGCAAG- ----GTGGGG 3101
Qy 1141 CAAGCNAGATGCTCTCCAGGATTTCAACCTGAGCCCTGCCANCCCTGCTGAADA 1200
Db 3102 GCAAGCAGATGCTCTCCAGGATTTTCAACC- --TGAGCCCTTGCCTCCACCTGCTGAGAA 3158
Qy 1201 ARAAYTNCCGCCAGCTGAGAGACACAGAGGATGNCAGGATTTNNACCTTGGGGAA 1260
Db 3159 AACACT- --CGCCACGTTGAAGAGACA- ---GAGAGGATGGCAGGAGTTACCTCGGGAAA 3212
Qy 1261 AAAAAACAGGAGTCTTTNTTTCTGCCCTGCTCCAGTNCAGTTCGAGTTGGCCTGNAACCCGCTTG 1320
Db 3213 CAAACAGGATCTTCTGCTCCCTGCTCCAG- ----TCCAGTTGGCCTGACCCGCTGG- 3264
Qy 1321 ANTCACTGACATTTGTTGGCAGCAGGGGAGCAGCTTCAGCTCGGTCAGAGGG 1380
Db 3265 -ATCAGTGACATTTGCTGGCAGA- CAGGGGAGAGCAGCTTCCAGCTCGGTCAGAGGG 3322
Qy 1381 GTGGGAGGCGCTTCGGCCCTCACCTNCCAGGCTGCTGNAGAGTGTCAAGTGTGA 1440
Db 3323 GTGGGAGGCGCTTCGGCCCTCACCT- CCAGCTGCTGTG- AGAGTGTCAAGTGTGA 3380
Qy 1441 AGGNCCTCAAACTCAGGNTTCAGTGCAGAACAGGTCNACAGGATATGCCCGCCCGNTA 1500

Db 3381 AGGGCCCCAACTC---AGGTCAGTGCGAACCAGGT-CAGCAGGATGCCCCGCCGTAG 3436
Qy 1501 GGTAAANNGGGGCGCTCTNAAACCCCTTGCTNGGCCCTNACCTNGGCCAGCTCANCCC 1560
Db 3437 GTTAA---GGGGGCGCTTAAACCCCTTGCTGCGCTCA---CCTGGCCAGCTCA-CCC 3488
Qy 1561 CTTTGGGTGTAGGGGAAAGAAATCCCTGAGACCCCTGGGAAGGCTWCCCTGGTAGAATACAC 1620
Db 3489 CTTTGGGTGTAGGGGAAAGAAATCCCTGAGACCCCTGGGAAGGCT-CCCTGGGTAGAATACAC 3547
Qy 1621 CACACTTTTCAGTTGTTGCAACACAGGTCCTGAGTTCAGCTCCTGCTTCCAGCCAGGACC 1680
Db 3548 CACACTTTTCAGTTGTTGCAACACAGGTCCTGAGTTCAGCTCCTGCTTCCAGCCAGGACC 3607
Qy 1681 AAGAAGGTGTAGTGAAGTGGTCTCAGTNCGCCAGACATGTCGCCCTTTGCTGCTG 1740
Db 3608 AAGAAGGTGTAGTGAAGTGGTCTCAGT-CCCCAGACATGTCGCCCTTTGCTGCTG 3666
Qy 1741 GCTACCACTCTTCCCAGAGCAGAGGCCCGAGCCCTTTCAGGCCAGCACTGCCCCAG 1800
Db 3667 GCTACCACTCTTCCCAGAGCAGAGGCCCGAGCCCTTTCAGGCCAGCACTGCCCCAG 3726
Qy 1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGTGAAGGTGATGCAAGGATATGCTGA 1860
Db 3727 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGTGAAGGTGATGCAAGGATATGCTGA 3786
Qy 1861 CAGGAACAGTCTGTGGATGGACATCATCAGTGTNCCAGCGTGTCTGCTGCTGCTGCTG 1920
Db 3787 CAGGAACAGTCTGTGGATGGACATCATCAGTGTCTAA--GGAAAGCAGCAGAGAGAGCGC 3844
Qy 1921 TCCGGCGCCAGNCCCACTNATCAGTGTNCCAGCGTGTCTGCTGCTGCTGCTGCTG 1980
Db 3845 TCCGGCGCCAGNCCCACTNATCAGTGT-CCAGCGTGTCTGCTGCTGCTGCTGCTG 3893
Qy 1981 GCTNAGNATCANCACTGACACTNCACTTNGCCCTGCTGCTGCTGCTGCTGCTGCTG 2040
Db 3894 GCACAGCTCAGCATCACATGACACTCACCCTGCTGCTGCTGCTGCTGCTGCTGCTG 3951
Qy 2041 CCGNACGACACTTTCACNTCTGATGACCTCAAGACCTTTCATGCTGCTGCTGCTGCTG 2100
Db 3952 CGG-ACGGACCTTTCGAC-TCGATG-ACCTCAAGACACTTTCATGCTGCTGCTGCTG 4008
Qy 2101 GCAGGNCAGGNCAGGNCAGTGACANCTGTAGNAGCATANGAANGCCAGGAGATGG 2160
Db 4009 GGGCAGGCVAGGCGAGTGACACTGT-AGGAGCATAGCAAGCCAGGAGATGG 4058
Qy 2161 GGTGNAAGGNCACAGCTCTGAGCTGTCCANCATGATGTGACATNCCCTCAAACTCTTN 2220
Db 4059 GGTGAA--GGGACACAGCTTGTAGCTGTCCA-CATGCACTGTGACT-CCTCAAACTCTT- 4113
Qy 2221 NCCAGNATTTCTCTAAGAAATAGCANCCTTNGCCCAATGCCCCAGCTTAGCCTCTTCT 2280
Db 4114 --CGAGATTTCTTAAGAAATAGCACCCTT-ATTCCCCATGCCCCAGCTTAGCCTCTTCT 4169
Qy 2281 CCCAGGGAGCTANCTCAGGACTCAGTAGCATTAATCAGCTGTGNAATCGTCAGGGGG 2340
Db 4170 CCCAGGGAGCTA-CTCAGGACTCAGTAGCATTAATCAGCTGTG-AATCGTCAGGGGG 4227
Qy 2341 TGTCTGTAGCTCAACCTCTTGGGGAGGGGAGCCCGAGACTCCGTGGGAGAGCTCAT 2400
Db 4228 TGTCTGTAGCTCAACCTCTTGGGGAGGGGAGCCCGAGACTCCGTGGGAGAGCTCAT 4287
Qy 2401 TCCACATCTGCCAAGACAGCTTTNGTTCAGCTGTCCCATTTGAGTCAGCTGCTCCC 2460
Db 4288 TCCACATCTTGCCAAGACAGCTTTT-GTTCAGCTGTCCCATTTGAGTCAGCTGCTCCC 4346
Qy 2461 GGGGAGAGAGCCCCCGCCCCAGCACATAAAGAACTGCAAGCTTGGTACTGCAAGTCTG 2520
Db 4347 GGGGAGAGAGCCCCCGCCCCAGCACATAAAGAACTGCAAGCTTGGTACTGCAAGTCTG 4406
Qy 2521 GGTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGTATGACAAATGTT 2572
Db 4407 GGTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGTATGACAAATGTT 4458

RESULT 5
US-60-659-397-566
; Sequence 566, Application US/60659397
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C001470
; CURRENT APPLICATION NUMBER: US/60/659,397
; CURRENT FILING DATE: 2005-03-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 566
; LENGTH: 4684
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-659-397-566

Query Match 67.7%; Score 1781.4; DB 15; Length 4684;
Best Local Similarity 89.1%; Pred.No. 0;
Matches 2292; Conservative 10; Mismatches 183; Indels 87; Gaps 35;

Qy 1 CCCCTCTCACAGCCAGGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCGCGTGT 60
Db 2150 CCCCTCTCACAGCCAGGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCGCGTGT 2209
Qy 61 CTGAGCGGAGCTCGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120
Db 2210 CTGAGCGGAGCTCGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 2269
Qy 121 GCCCTTGGAGGGGAGAATATAAGAAACAAGACATCCACCGCCAAATCAAGCCCAATTACC 180
Db 2270 GCCCTTGGAGGGGAGAATATAAGAAACAAGACATCCACCGCCAAATCAAGCCCAATTACC 2329
Qy 181 ACCAGACCTCCATGCGCCAGCGAGAGCTTTCGCAAGGGCCCCAGGGCCCCCGGCGAG 240
Db 2330 ACCAGACCTCCATGCGCCAGCGAGAGCTTTCGCAAGGGCCCCAGGGCCCCCGGCGAG 2389
Qy 241 CTGAGGAGACAAAGGAGAGCCCTTAAGCTTCCAGCCTCTCTCCACAGAGCCCCCAG 300
Db 2390 CTGAGGAGACAAAGGAGAGCCCTTAAGCTTCCAGCCTCTCTCTCCACAGAGCCCCCAG 2449
Qy 301 AGCCAAACAAGTCTCTCTTGAATTGTGACAAAGGAGAGTCTGGGATGTGGAAACCTT 360
Db 2450 AGCCAAACAAGTCTCTCTTGAATTGTGACAAAGGAGAGTCTGGGATGTGGAAACCTT 2509
Qy 361 TACCTCTGTCTCTCTGGAGCCAGCCCTGCGCAAGAAACCCAGCTCAACAGCGGGAAG 420
Db 2510 TACCTCTGTCTCTCTGGAGCCAGCCCTGCGCAAGAAACCCAGCTCAACAGCGGGAAG 2569
Qy 421 CAACCGTCCCAGGAGCAGGAACCTGAGCAGCTGGAATAGAAATATTCTCAACAGCCTGT 480
Db 2570 CAACCGTCCCAGGAGCAGGAACCTGAGCAGCTGGAATAGAAATATTCTCAACAGCCTGT 2629
Qy 481 CCCAGCAATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATPCACAGCC 540
Db 2630 CCCAGCAATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATPCACAGCC 2689
Qy 541 TCTCCCTGTGGATGACAGTGAGAGAACCCATCAAGGCCCTCTCAAGCTCGCGGGACA 600
Db 2690 TCTCCCTGTGGATGACAGTGAGAGAACCCATCAAGGCCCTCTCAAGCTCGCGGGACA 2749
Qy 601 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGAGGCTCGAAGCTCCAGCTGGA 660
Db 2750 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGAGGCTCGAAGCTCCAGCTGGA 2809
Qy 661 ACATGGTCTGGCCCCGGGGCGGCCACCGACACCCCAAGCTATTTCATATGGTGTGAAAG 720

Db 2810 ACATGGTGTGGCGGGGGGGCCACCCGACACCCCAAGCTATTTCAATGGTGTGAAG 2869
Qy 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGCAATCCGGAGTTCCACCGGTCAAAG 780
Db 2870 TCCAAATACAGTCTCTTAATGGTGAACACCTGCAATCCGGAGTTCCACCGGTCAAAG 2929
Qy 781 TGGGAGACATCGGCACCTGGCATCAGCAGCCAGATCCAGCTGCAGCCCTTCAGCTTGTCA 840
Db 2930 TGGGAGACATCGGCACCTGGCATCAGCAGCCAGATCCAGCTGCAGCCCTTCAGCTTGTSA 2989
Qy 841 CCAAAGACGGGACGCTGTTGCTTACGATCAGATGGAGGTGCCAGACTCCGGGATCGACCTGC 900
Db 2990 CCAAAGACGGGACGCTGTTGCTTACGATCAGATGGAGGTGCCAGACTCCGGGATCGACCTGC 3049
Qy 901 AGTGCACACTGGCCCCCTGATGGCAGCTTGGCTGGAGCTGGAGGTCAAGCATGGCCAGC 960
Db 3050 AGTGCACACTGGCCCCCTGATGGCAGCTTGGCTGGAGCTGGAGGTCAAGCATGGCCAGC 3109
Qy 961 TGGAGAACAGCCCTAACCCCTGCCCTCCACCGCCGGCTCCACACTGCGGGAAGCAGCCT 1020
Db 3110 TGGAGAACAGCCCTAACCCCTGCCCTCCACCGCCGGCTCCACACTGCGCGG-AAGCAGCCT 3168
Qy 1021 TCCCTGCTGGTGCACGATGCTGCCCTGAAACACAGCTCAGCCGTTTCCAGGGGATVTG 1080
Db 3169 TCCCTGCTGGTGCACGATGCTGCCCTG-AAACACAGCTCAGCCGTTTMCAGGG---ATC 3224
Qy 1081 NCCAGCCCCCGGCTCAGCAGNTGGGAACACGAGGCCCTCGNCAGNAGCNAAGTNGGGG 1140
Db 3225 TGCAGCCCCCGGCTCAGCAGTGGGACAGGGCCCTCGCAGCAGCAAG-----GTGGG 3277
Qy 1141 CAAGCNAAGATGCTCCAGGATTTTCAACCTGAGCCNTGCCCCANCCCTGCTGAADA 1200
Db 3278 GCAAGCAGAAATGCTCCAGGATTTTCAACC---TGAGCCCTGCCCCACCCCTGCTGAGAA 3334
Qy 1201 AACAAYTNCGCCACGCTGAGAGACAGAGAGATGNCAGGATTTNACCTTGGGAA 1260
Db 3335 AACAAT---CGGCCACGTTGAAGAGACA---GAGGAGATGGCAGGATTTACCTCGGAAA 3388
Qy 1261 ACAAACAGGAGATCTTTTTCGCCCTGCTCCAGTNCAGTGGCCCTGNAACCCGCTGG 1320
Db 3389 CAACAGGATCTTCTGCCCTGCTCCAG-----TGAGTTGGCTGACCCGCTGG- 3440
Qy 1321 ANTCAGTGACATTTGTCAGANAGGGAGAGAGCTTCCAGCTGGGTGAGAAGG 1380
Db 3441 -ATCAGTGAACATTTGCTGGCAGA-CAGGGAGAGCAGCTTCCAGCTGGGTGAGAAGG 3498
Qy 1381 GTGGGAGAGCCCTTCGGCCCTCACCCTNCAGGCTGCTGTGNAGAGTGTCAAGTGTGA 1440
Db 3499 GTGGGAGAGCCCTTCGGCCCTCACCCT-CCAGGCTGCTGTG-AGAGTGTCAAGTGTGA 3556
Qy 1441 AGGNCCTCAANCTCAGNTTTCAGTGAGAACACAGGTNCAGCAGGTATGCCCGCCGNTA 1500
Db 3557 AGGCCCTCAATC---AGGTTTCAGTGAGAACACAGGT-CAGCAGGTATGCCCGCCGCTAG 3612
Qy 1501 GGTTAANNNGGGGCCCTCTNAACCCCTTGCTNGGCTNCACCTNGGCCAGCTCANCCC 1560
Db 3613 GTTAA---GGGGCCCTCTAAACCCCTTGCTGCGCTCA---CCTGGCCAGCTCA-CCC 3664
Qy 1561 CTTTGGGTGAGGGGAAAAGATTCCTGACCCCTGGGAAGGCTWCCCTGGTAGAATACAC 1620
Db 3665 CTTTGGGTGAGGGGAAAAGATTCCTGACCCCTGGGAAGGCT-CCCTGGTAGAATACAC 3723
Qy 1621 CACATTTTTCAGGTGTTGCAACAGGCTCTGAGTTGACCTCTGTTTCAGCCAGGACC 1680
Db 3724 CACATTTTTCAGGTGTTGCAACAGGCTCTGAGTTGACCTCTGTTTCAGCCAGGACC 3783
Qy 1681 AAAGAAGGTGTGAAGTGAAGTGTCTCAGTNCACCCAGACATGTCGCCCTTTGCTGCTG 1740
Db 3784 AAAGAAGGTGTGAAGTGAAGTGTCTCAGT-CCCAGACATGTCGCCCTTTGCTGCTG 3842
Qy 1741 GCTACCACTCTTCCCCAGAGCAGCAGGCCCGGAGCCCTTTCAGGCCAGCACTGCCCCAG 1800
Db 3843 GCTACCACTCTTCCCCAGAGCAGCAGGCCCGGAGCCCTTTCAGGCCAGCACTGCCCCAG 3902

Qy 1801 ACTCGCTGGCAGTCACTGTTCCCTCATCTGTAAAGGTGAAGGGTATGACAGATATGCCTGA 1860
Db 3903 ACTCGCTGGCAGTCACTGTTCCCTCATCTGTAAAGGTGAAGGGTATGACAGATATGCCTGA 3962
Qy 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGAGAGY 1920
Db 3963 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGAGAGC 4020
Qy 1921 TCCGGGGCCCCAGNCCCACACTNATCAGTGTNCCAGCGTGTCTNGGTTNCCCAGNAGCACA 1980
Db 4021 TCCGGGGCCCCAGNCCCACACTNATCAGTGT---CCAGCGTGTCTGTTCCC-----CAGA 4069
Qy 1981 GCTNAGNCAATCANCACTGACATNCACTTNGCCCTGCCCTTNGGCCANGAGGGTACTG 2040
Db 4070 GCACAGCTCAGCATCACTGACACTCACCTGCTGCCCTGCCCTGCG---CCAGAGGGTACTG 4127
Qy 2041 CCGNAGCGCACTTTCGACNTCTGATGNACCTCAAGACATTTTCATGCTGCTNGCCCTCTNNG 2100
Db 4128 CCG-ACGGCACTTTGAC-TCTGATG-ACCTCAAGACATTTTCATGCTGCTGCCCTCTGGCA 4184
Qy 2101 GCAGGNCAGGNCAGGNCAGTGCANCTGTAGNAGCATANGCAANGCCAGGAGATGG 2160
Db 4185 GGGCAGGGYAGGAGTGCACACTGT-----AGGACATAGCAAGCCAGGAGATGG 4234
Qy 2161 GGTGNAAGGANCACTGTTGAGCTGTCCANCAATGCAATGCAATGCTCAAACTCTTN 2220
Db 4235 GGTGAA--GGGACACAGTCTGAGCTGTCCA-CATGATGTGACT-CCTCAAACTCTT- 4289
Qy 2221 NCCAGNATTTCTTAAGATAGCANCCCCCTTNGCCCTTNGCCCTGAGCTTACCTCTTCT 2280
Db 4290 --CCAGATTTCTTAAGATAGCACCCCC--TTTCCCATTTGCCCCAGCTTAGCCTCTTCT 4345
Qy 2281 CCAGGGGAGCTANCTCAGGACTCACGTAGCATTAATCAGCTGTGNAATCGTCAGGGG 2340
Db 4346 CCAGGGGAGCTA-CTCAGGACTCAGTAGCATTAATCAGCTGTG-AATCGTCAGGGG 4403
Qy 2341 TGTCTGTAGCTCAACTCTTGGGGCAGGGGACGCCAGACTCCGTTGGGAGAAAGCTCAT 2400
Db 4404 TGTCTGTAGCTCAACTCTTGGGGCAGGGGACGCCAGACTCCRTGGGAGAAAGCTCAT 4463
Qy 2401 TCCACATCTTGCAGAGAGAGCTTTNGTCCAGCTGTCCACATTTGAGTCAGCTGCTCCC 2460
Db 4464 TCCACATCTTGCAGAGAGAGCTTTT-GTCCAGCTRTCCACATTTGAGTCAGACTGCTCCC 4522
Qy 2461 GGGGAGAGAGCCCCCGGCCCCAGCACATAAAGAACTGCAGCCTTGTGTAAGTGTG 2520
Db 4523 GGGGAGAGAGCCCCCGGCCCCAGCACATAAAGAACTGCAGCCTTGTGTAAGTGTG 4582
Qy 2521 GGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTGGGGTGTGACAAATGTT 2572
Db 4583 GGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTGGGGTGTGACAAATGTT 4634

RESULT 6

US-60-659-397-12129
; Sequence 12129, Application US/60659397
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/659,397
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12129
; LENGTH: 65967
; TYPE: DNA
; ORGANISM: Homo sapiens

US-60-659-397-12129

Query Match 36.9%; Score 971.8; DB 15; Length 65967;
Best Local Similarity 84.2%; Pred. No. 1.6e-222; Mismatches 183; Indels 87; Gaps 35;
Matches 1483; Conservative 9;
QY 811 AGATCCAGCTCAGCGCTTACGTTGGTGCACAAAGAGGGGAGCGCTGTTCGCTACGACA 870
DB 58244 AGATCCAGCTCAGCGCTTACGTTGGTGCACAAAGAGGGGAGCGCTGTTCGCTACGACA 58303
QY 871 TGGAGGTGCAGACTCGGCGATCGACTGCGATGACACCTGCGCCCTCTGATGGCAGCTTCG 930
DB 58304 TGGAGGTGCAGACTCGGCGATCGACTGCGATGACACCTGCGCCCTCTGATGGCAGCTTCG 58363
QY 931 CCTGAGCTGGAAGGTCAAGCATGCCAGCTGAGCAAGAGCCCTAAACCTGCCCTCCAC 990
DB 58364 CCTGAGCTGGAAGGTCAAGCATGCCAGCTGAGCAAGAGCCCTAAACCTGCCCTCCAC 58423
QY 991 CGCCGGCTCCACACTCGCGGAAAGCAGCCTTCTGCTCGGTGCACGATGCTGCCCTGAAA 1050
DB 58424 CGCCGGCTCCACACTCGCGG--AAGCAGCCTTCTGCTCGGTGCACGATGCTGCCCTG-AA 58481
QY 1051 ACACAGGTTCAGCGCTTCCAGGGGATYTNCCAGCCCGCCCGCTTCACAGNVTGGGAACC 1110
DB 58482 ACACAGGTTCAGCGCTTCCAGGGG---ATCTGCCAGCCCGCCCGCTTCAGCAGTGGGACCCAG 58538
QY 1111 AGGGCTCGNACAGNAGNAGGTNGGGGCAAGCAGNAGATGCCCTCCAGGATTTTCACAN 1170
DB 58539 GGCCTTCGACAGCAAG-----GTGGGGGCAAGCAGATGCCCTCCAGGATTTTCACAC 58591
QY 1171 CCTGAGCCCTCCACCTCGCTGGAADAAAACATNCCGCGCAGCTGAAAGAGACAGAAG 1230
DB 58592 C---TGAGCCCTGCCCGCCCTGCTGAGAAACACT--CCGCGCAGTGAAGAGACA---- 58642
QY 1231 GAGGATGNCAGGATGTTNACCTYGGGGAACAAACAGGGATCTTTNTTCTGCCCTGTC 1290
DB 58643 GAGGAGGATGCGAGGATTTACCTCGGGAACAAACAGGATCTTCTGCTGCCCTGCTCCAG- 58701
QY 1291 TCCAGTNCAGTTCGCTGNACCCGCTTGGANTCAGTACCACTTTTGGGAGNACAGGG 1350
DB 58702 -----TCGAGTTGGCCCTGA---CCGCTTGGATCAGTACCACTTTGCTGGCAGA--CAGGG 58752
QY 1351 GAGAGCAGCTTCCAGCCTGGGTGAGAAAGGGGTGGCGAGCCCTTCGGGCCCTCACCTTNC 1410
DB 58753 GAGAGCAGCTTCCAGCCTGGGTGAGAAAGGGGTGGCGAGCCCTTCGGGCCCTCACCTT-C 58811
QY 1411 CAGGCTGCTGNAGAGTGTCAAGTGTGTAAGGNCNCCAAANCTCAGGNTTCAGTGCAGA 1470
DB 58812 CAGGCTGCTGTG-AGAGTGTCAAGTGTGTAAGGGCCCAAACTC---AGGTTTCAGTGCAGA 58867
QY 1471 ACCAGGTNCAGCAGGTATGCCCGCCGNTAGGTTAANNNGGGGCCCTCTNAAACCCCTTG 1530
DB 58868 ACCAGGT-CAGCAGGTATGCCCGCCGNTAGGTTAA-----GGGGGCCCTCTAAACCCCTTG 58922
QY 1531 CCTNGGCCCTNCACTTNGGCCACTCANCCCTTTTGGGTGTAGGGGAAAGAAATGCCTGA 1590
DB 58923 CTTGACCTCA---CCTGGCCAGCTCA-CCTCTTTGGGTGTAGGGGAAAGAAATGCCTGA 58978
QY 1591 CCTCGGAAGGTTCCTGGTGAATATACACACATTTTCAGTGTGTGCAACACAGGTC 1650
DB 58979 CCTCGGAAGGT-CCCTGGTGAATATACACACATTTTCAGTGTGTGCAACACAGGTC 59037
QY 1651 CTGAGTTCACCTCTGTTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGTTCCTCA 1710
DB 59038 CTGAGTTCACCTCTGTTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGTTCCTCA 59097
QY 1711 GTNCCCAGACATGTGCCCTTTGCTGCTGGCTACCACTCTTCCCAGAGCAGAGGCC 1770
DB 59098 GT-CCCCAGACATGTGCCCTTTGCTGCTGGCTACCACTCTTCCCAGAGCAGAGGCC 59156
QY 1771 CGAGGCCCTTCAGGCCCAGCACTGCCCGNAGCTCGCTGGCACTCAGTTCCTCATCTGTA 1830
DB 59157 CGAGGCCCTTCAGGCCCAGCACTGCCCGNAGCTCGCTGGCACTCAGTTCCTCATCTGTA 59216

QY 1831 AAGGTGAAGGCTGATGACGATATGCTCTGACAGGAACAGTCTGTGGATGGACATGATCAG 1890
DB 59217 AAGGTGAAGGCTGATGACGATATGCTCTGACAGGAACAGTCTGTGGATGGACATGATCAG 59276
QY 1891 TGCTNAAGNAAAGCAGCAGAGAGAGAGCGYTCGCGCGCCCGCCAGNCCCACATNATCAGTGT 1950
DB 59277 TGCTAA--GGAAAGCAGCAGAGAGAGAGCGTCCGCGCGCCCGCCAGCCACTATCAGTGT-- 59332
QY 1951 NCCAGCTGCTNGGTTNCCCAGNAGCAGACACTNCACTNCACTGACACTNCACTCACC 2010
DB 59333 --CAGCGCTGCTGGTTCCTCC-----CAGAGCAGACTCAGCATCAGACTGACACTCACC 59383
QY 2011 TNGCCCTGCCCTTNGGCCANGAGGCTACTGCCGNAAGGCACTTTTGACANTCTGTAGTGNACC 2070
DB 59384 CTGCCCTGCCCTCG--CCAGAGGGTACTCGG-ACGGCACTTTTGAC-CTGTATG-ACC 59438
QY 2071 TCAAAGCACTTTTCATGCTNGCCCTCTNNCGCAGGNCAGGNCAGGNCAGTGAACNCT 2130
DB 59439 TCAAAGCACTTTTCATGCTGCTGCCCTCTGCGAGGGCAGGGYAGGGCAGTGACACT----- 59491
QY 2131 GTAGNAGCATANGCAANGCCAGGAGATGGGGTGNAGGGGANCACAGTCTTTGAGCTGTCC 2190
DB 59492 ---GTAGGAGCATAGCAAGCCAGGAGATGGGGTGAA--GGGACACAGTCTTTGAGCTGTCC 59546
QY 2191 ANCATGCACTGTGACTNCTCAAACCTCTNNCCAGNATTTCTTAAGAAATAGCANCCCC 2250
DB 59547 A-CATGCACTGTGACT-CCTCAAACTCTT---CCAGATTTCTTAAGAAATAGCACCCCC- 59600
QY 2251 TTNCCCATTGCCCAGCTTACGCTCTTCTCCCAGGGAGCTANCTCAGGACTCAGCTAG 2310
DB 59601 -TTCCCATTGCCCAGCTTACGCTCTTCTCCCAGGGAGCTA-CTCAGGACTCAGCTAG 59658
QY 2311 CATTAAATCAGCTGTGNATCGTCAGGGGTGTCTGTAGGCTCAACCTCTGGGGCAGG 2370
DB 59659 CATTAAATCAGCTGTG--AATCGTCAGGGGTGTCTGTAGGCTCAACCTCTGGGGCAGG 59717
QY 2371 GGAGCCGAGACTCCGTGGGAGAAAGTCAATTCACACATCTTGCACAGACAGCTTTTNGTC 2430
DB 59718 GGAGCCGAGACTCCTGGGAGAAAGTCAATTCACACATCTTGCACAGACAGCTTT-CTC 59776
QY 2431 CAGCTGTCCAATTCAGTTCAGACTGCTCCCGGGAGAGAGCCCGGGCCCCCAGACACATAA 2490
DB 59777 CAGCTGTCCAATTCAGTTCAGACTGCTCCCGGGAGAGAGCCCGGGCCCCCAGACACATAA 59836
QY 2491 AGAAGTSCAGCTTGGTACTGACAGAGTCTGGGTGTGAGAGAACTTTTGTAGCAATAA 2550
DB 59837 AGAAGTSCAGCTTGGTACTGACAGAGTCTGGGTGTGAGAGAACTTTTGTAGCAATAA 59896
QY 2551 GTTTGGGTGATGACAAATGTT 2572
DB 59897 GTTTGGGTGATGACAAATGTT 59918

RESULT 7

US-11-060-756-560
; Sequence 560: Application US/11060756
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 560
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-560

Query Match	29.2%;	Score 769.2;	DB 13;	Length 1400;
Best Local Similarity	82.9%;	Pred. No. 2.6e-174;		
Matches 1224;	Conservative	8;	Mismatches 166;	Indels 79; Gaps 31;
Qy	1103	TGGGAACACAGGCGCTCGNCACGACGNAAGGTGNGGGGCGAAGCAGNAGATGCCTCCACGGA	1162	
Db	3	TGGGAACACAGGCGCTCG-----CAGCAGCAAGGTGGGGGCAAGCAGAAATGCCTCCACGGA	57	
Qy	1163	TTTCACANCTGAGCCCNVTGCCCANCCCTGCTCTGAADAAAACAATNCCGGCCACGTGAAGA	1222	
Db	58	TTTTCACACC---TGAGCCCTGCCCCACCCCTGCTGARAANAACACT--CGCCACGTGAAGA	112	
Qy	1223	GACAGAAGGAGGATGNCNACGAGGTNNACCTYGGGGAAAACAAAACAGGATCTTTNTTCT	1282	
Db	113	GACAGAAGGAGGATGGCAGGAGTT---ACCTSGGGAAACAAACAGGATCTTCTGCGCCT	169	
Qy	1283	GCCCTGCTCCAGTNCAGTTCAGTTCGGCTGNAACCGCTTCGANTCAGTGCACATTTGTTGGCA	1342	
Db	170	GCTCCAG-----TCGAGTTGGCTCTGACCCGCTTGG--ATCAGTGACCAATTTGTYGGCA	220	
Qy	1343	GANCAGGGGAGAGCAGCTTCCAGCGCTGGGTGAGAAGGGGTGGGCGAGCCCTTCGGGCCCT	1402	
Db	221	GA-CAGGGGAGAGCAGCTTCCAGCCTGGGTGAGAAGGGGTGGGCGAGCCCTTCGGGCCCT	279	
Qy	1403	CACCTTNCAGGCTGCTGTGNAGAGTGTCAAGTGTGAAGGNCACAAAACCTCAGNNTC	1462	
Db	280	CACCTT--CCAGGCTGCTGTG--AGAGTGTCAAGTGTGAAGGGCCCAACTC---AGGTTTC	334	
Qy	1463	AGTGCAGAACGAGTNCAGCAGGTATGCCCGCCGNTAGGTTAANNNGGGGGCCCTCTNAA	1522	
Db	335	AGTGCAGAACGAGGT--CAGCAGGTATGCCCGCCGTAAGTTAA----GGGGGGCCCTCTAA	389	
Qy	1523	ACCCCTTGGCTTNGCCCTNCACCTTNGGCGAGCTCANCCCTTTTGGGTGTAGGGGAAAAGA	1582	
Db	390	ACCCCTTGGCTTGGCTCA---CCTGGCGAGCTCA--CCCTTTTGGGTGTAGGGGAAAAGA	445	
Qy	1583	ATGCTGACCTTGGAAAGGCTWCCTGGTGAATACACACATTTTTCAGGTTGTTGCAA	1642	
Db	446	ATGCTTGACCTTGGAAAGGCT--CCCTGTTGAATACACACATTTTTCAGGTTGTTGCAA	504	
Qy	1643	CACAGTTCCTGAGTTGACCTCTGGTTTCAGCCNAGGACCAAGAAGGTGTGTAAAGT	1702	
Db	505	CACAGTTCCTGAGTTGACCTCTGGTTTCAGCCNAGGACCAAGAAGGTGTGTAAAGT	564	
Qy	1703	GGTTCTCAGTNCACACATGTGCCCTTTGCTGCTGGCTACACACTTCTCCCCAGACGA	1762	
Db	565	GGTTCTCAGT--CCCCAGACATGTGCCCTTTGCTGCTGGCTACCACTTCTCCCCAGACGA	623	
Qy	1763	GCAGGCCCCGAGGCCCTTTCAGGCCCCAGCACTGCCCCAGATCTGCTGTCAGTTCCTCT	1822	
Db	624	GCAGGCCCCGAGGCCCTTTCAGGCCCCAGCACTGCCCCAGATCTGCTGTCAGTTCCTCT	683	
Qy	1823	CATCTGTAAAGGTGAAGGTTATGTCAGGATATGCTCTGACAGGAACAGTCTGTGGATGGAC	1882	
Db	684	CATCTGTAAAGGTGAAGGTTATGTCAGGATATGCTCTGACAGGAACAGTCTGTGGATGGAC	743	
Qy	1883	ATGATCAGTCTTAAAGGNAAGCAGCAGAGAGACGTYCTCGGCGCCCCCAGNCCCACCTN	1942	
Db	744	ATGATCAGTCTTAA--GGAAAGCAGCAGAGAGACGCTCTCGGCCCCCAGCCTCATAT	801	
Qy	1943	ATCAGTGTNCAGCGTCTGTTGTTNCCCCCAGNAGCACAGCTNCAGNATCANCACTGACA	2002	
Db	802	CAGTGT---CCAGCGTCTGCTTCCC-----CAGAGCACAGCTCAGCATCACACTGA	850	
Qy	2003	CTNCACCTTNGCCCTTGGCCCTTNGCCANGAGGTTACTGCGNAGCGCACTTTGCACTCT	2062	
Db	851	CATCTACCTTGGCCCTTGGCCCTTGG---CCAGAGGTTACTGCGC--ACGCGCACTTTGCACT	905	
Qy	2063	GATGNACCTCAAAGCACTTTTCATGCGTNGCCCTCTNNNGCAGGNCAGGNCAGGNCAG	2122	
Db	906	TGATGACCTCAAAGCACTTTTCATGCGTNGCCCTCTTGGCAGGGGAGGCGAGGTCAGAC	965	
Qy	2123	TGCANCTGTAGGNAGCATANGCAANGCCAGGAGATGGGGTGTGNAAGGANCACAGTCTTG	2182	

Db	966	T-----GTAGGAGCATATGCAAGCCAGGAGATGGGGTGAA--GGGACACAGTCTTTG	1013
Qy	2183	AGCTGTCCANCATGCATGTCACTNCCCTCAAACCCTCTTNCCAGNATTTCTCTAAGAATAG	2242
Db	1014	AGCTGTCCA-CATGATGTGACT-CTTCAAACCTCTT---CCAGATTTCCTAAGAATAG	1068
Qy	2243	CANCCCCCTTNNCCCCCATTGCCCGCAGCTTAGCCCTCTTCTCCCAGGGGAGCTANCTCAGGAC	2302
Db	1069	CACCCCC--TTCCCCATTGCCCGCAGCTTAGCCCTCTTCTCCCAGGGGAGCTA-CTCAGGAC	1125
Qy	2303	TCACGTAGCATTTAAATCAGCTGTGNAAATCGTCAAGGGGTGTCTGTAGCTTCAACCTCCT	2362
Db	1126	TCACGTAGCATTTAAATCAGCTGTG-AATCGTCAAGGGGTGTCTGTAGCTTCAACCTCCT	1184
Qy	2363	GGGSCAGGGGACGCCGAGACTCCGCTGGGAAGCTCATTTCCCACTCTTGCCTCAAGACAGC	2422
Db	1185	GGGSCAGGGGACGCCGAGACTCCGCTGGGAAGCTCATTTCCCACTCTTGCCTCAAGACAGC	1244
Qy	2423	CTTTNGTCCAGCTGTCCACAATTGAGTCAGACTCTCTCCCGGGAGAGAGCCCCCGCCCCCA	2482
Db	1245	CTTT-GTCCAGCTGTCCACAATTGAGTCAGACTCTCTCCCGGGAGAGAGCCCCCGCCCCCA	1303
Qy	2483	GCACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGGTGTGTAGAGAACTCTTTGTAA	2542
Db	1304	GCACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGGTGTGTAGAGAACTCTTTGTAA	1363
Qy	2543	GCAATTAAGTTTCGGGTGATGACAAATGTTAAAAAAA	2579
Db	1364	GCAATTAAGTTTCGGGTGATGACAAATGTTAAAAAAA	1400
 RESULT 8 US-11-060-756-4832 ; Sequence 4832, Application US/11060756 ; GENERAL INFORMATION: ; APPLICANT: Wyeth ; APPLICANT: Mounts, William Martin ; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles o ; TITLE OF INVENTION: Target Genes ; FILE REFERENCE: AM101083 (031896-042000) ; CURRENT APPLICATION NUMBER: US/11/060,756 ; CURRENT FILING DATE: 2005-02-18 ; NUMBER OF SEQ ID NOS: 303284 ; SOFTWARE: PatentIn version 3.2 ; SEQ ID NO 4832 ; LENGTH: 1400 ; TYPE: DNA ; ORGANISM: Homo sapiens US-11-060-756-4832			
 Query Match 29.2%; Score 769.2; DB 13; Length 1400; Best Local Similarity 82.9%; Pred. No. 2.6e-174; Matches 1224; Conservative 8; Mismatches 166; Indels 79; Gaps 31;			
Qy	1103	TGGGAACACGGGCTCTGNCAGCNAGNAGAGTNGGGGCAAGCAGAAATGCCTCCACAGGA	1162
Db	3	TGGGAACACGGGCTCTG-----CAGCAGCAGGTGGGGGCACGAGATGCCTCCACAGGA	57
Qy	1163	TTTTCAANCCTGAGCCCTGCCCCCANCCCTGCTGAADAAAAACAAYTNCGCCACGCTGAAGA	1222
Db	58	TTTTCACACC---TGAGCCCTTGCCCACTCTGCTGAPAAAAACACT--CCGCCACGCTGAAGA	112
Qy	1223	GACAGAGGAGGATGNCAGGAGTTNNACTCTGGGGGAACAAACAGGATCTTTNTTCT	1282
Db	113	GACAGAGGAGGATGGCAGAGT---ACCTSGGGAACAAACAGGATCTTYTCTGCCCT	169
Qy	1283	GCCCTGTCTCCAGTNCAGTTGGCCTGNACCCGCTTGGANTCAGTCACCATTTGTTGGCA	1342
Db	170	GCTCCAG-----TCAGTTGSCCTGACCCGCTTGG--ATCAGTCACCATTTGTTGGCA	220
Qy	1343	GANCAGGGGAGACAGCTTCCAGCTTGGGTGCAAGAGGGGTGGGCGAGCCCTTCGGCCCT	1402

Db 221 GA-CAGGGGAGACAGCTTCCAGCCTGGGTGTCAGAAAGGGGTGGGCGAGCCCTTCGGCCCT 279

Qy 1403 CACCCTNCCAGGCTGCTGAGAGCTGTCAGGTGTAAGGNCCTCAACATCTCAGNTTC 1462

Db 280 CACCTT-CAGGCTGCTGTG-AGAGTGTCAAGTGTGTAAGGCCCAACTC---AGGTTTC 334

Qy 1463 AGTGAGAACACAGGTCAGAGGATGATGCCCGCCGNTAGGTTAAANNNGGGGCCCTCTTAA 1522

Db 335 AGTGAGAACACAGGT-CAGCAGGTATGCCCGCCGCTAGGTTAA---GGGGGCCCTCTTAA 389

Qy 1523 ACCCTTGCCTNNGCTNACCTNNGCCAGCTCANCCCTTTTGGGTGTAGGGGAAAAGA 1582

Db 390 ACCCTTGCCTGGCTCA---CCTGGCCAGCTCA-CCTCTTTGGGTGTAGGGGAAAAGA 445

Qy 1583 ATGCTGACCCCTGGGAGGCTGCCCTGGTGTAGTAACACACACTTTTTCAGTGTGTGCAA 1642

Db 446 ATGCTGACCCCTGGGAGGCT-CCTGTGTAGTAACACACACTTTTTCAGTGTGTGCAA 504

Qy 1643 CACAGGTCTCTGAGTTGACCTCTGTGTTACGCAAGGACCAAGAGGCTGTAAAGTGAAGT 1702

Db 505 CACAGGTCTCTGAGTTGACCTCTGTGTTACGCAAGGACCAAGAGGCTGTAAAGTGAAGT 564

Qy 1703 GGTCTCAGTNCCTCAGACATGTGCCCTTTGCTGTGCTGTACACTCTTCCCAAGACA 1762

Db 565 GGTCTCAGT-CCCCAGACATGTGCCCTTTGCTGTGCTGTACACTCTTCCCAAGACA 623

Qy 1763 GCAGGCCCGAGCCCTTTCAGGCCAGACACTGCCCGCAGACTCGCTGGCACTCAGTTCCT 1822

Db 624 GCAGGCCCGAGCCCTTTCAGGCCAGACACTGCCCGCAGACTCGCTGGCACTCAGTTCCT 683

Qy 1823 CATCTGTAAAGGTGAAGGTGATGAGGATATGCTGTGACAGAAAGCTGTGTGATGAGAC 1882

Db 684 CATCTGTAAAGGTGAAGGTGATGAGGATATGCTGTGACAGAAAGCTGTGTGATGAGAC 743

Qy 1883 ATGATCAGTGTCTNAGGNAAGCAGACAGAGAGAGCGTTCGGGCCCCCAAGNCCCACTN 1942

Db 744 ATGATCAGTGTCTAA--GGAAAGCAGACAGAGAGAGCGTTCGGGCCCCCAAGNCCCACTAT 801

Qy 1943 ATCAGTGTNCCAGCGTGTCTGTTNCCCGCAGNAGCAGACTNCAGNATCANCATGACA 2002

Db 802 CAGTGT---CCAGCTGTGTGTCTCC-----CAGAGCAAGCTCAGCATCAGCTGA 850

Qy 2003 CTNACCTTNGCCCTGCCCCCTNNGCCANGAGGCTACTGCCGNACGGCACTTTGCACNTCT 2062

Db 851 CACTCACCTTGCCTGCCCCCTGG--CCAGAGGCTACTGCCG-ACGGCACTTTGCACT--C 905

Qy 2063 GATNACCTCAAAGCACTTTATGCTNGCCCTCTNNGGCGAGGNCAGGNCAGGNCAG 2122

Db 906 TGATGACCTCAAAGCACTTTATGCTGCTGCTCTGCGAGGGCAGGCGAGGTGACAC 965

Qy 2123 TGACANCTGTAGNAGCATANGCAANGCCAGGAGATGGGTGNAAGGANCACAGCTTTG 2182

Db 966 T-----GTAAGAGCATAGCAAGCCAGGAGATGGGGTAA--GGGACACAGCTCTTG 1013

Qy 2183 AGCTGTCCANCATGATGATGATNCTCAAACCTCTTNNCCAGNATTTCTCTAAGAAATAG 2242

Db 1014 AGCTGTCCA-CATGCAATGTGACT-CCTCAAACCTCTT---CCAGATTTCTCTAAGAAATAG 1068

Qy 2243 CANCCCCCTTNGCCCATTTGCCCCAGCTTAGCCCTCTTCCCGAGGAGGAGCTANTCTCAGAC 2302

Db 1069 CACCCCC--TTCCCCATTTGCCCCAGCTTAGCCCTCTTCTCCAGGGGAGCTA-CTCAGGAC 1125

Qy 2303 TCACCTAGCATTAATCAGCTGTGNAATCGTCAGGGGTGTCTGTAGCTCAACCTCCT 2362

Db 1126 TCACCTAGCATTAATCAGCTGTG-AATCGTCAGGGGGTGTCTGTAGCTCAACCTCCT 1184

Qy 2363 GGGGAGGGGAGCGCCAGACTCCGTGGGAGAGCTCATTTCCCACTTTTGCCAAGACAGC 2422

Db 1185 GGGGAGGGGAGCGCCAGACTCCGTGGGAGAGCTCATTTCCCACTTTTGCCAAGACAGC 1244

Qy 2423 CTTTNGTTCAGCTGTCCACATTTGAGTTCAGACTGCTCCCGGGGAGAGAGCCCGGGCCCCCA 2482

Db 1245 CTTT-GTCCAGCTGTCCACATTTGAGTTCAGACTGCTCCCGGGGAGAGAGCCCGGGCCCCCA 1303

Qy 2483 GCACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTGTAA 2542

Db 1304 GCACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTGTAA 1363

Qy 2543 GCAATAAAGTTTGGGTGATGACAAATGTTAAAAAAA 2579

Db 1364 GCAATAAAGTTTGGGTGATGACAAATGTTAAAAAAA 1400

RESULT 9

US-10-450-763-15295/c

; Sequence 15295, Application US/10450763

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

; FILE REFERENCE: 790CIP3/US

; CURRENT APPLICATION NUMBER: US/10/450,763

; CURRENT FILING DATE: 2003-06-11

; PRIOR APPLICATION NUMBER: PCT/US01/08631

; PRIOR FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: 09/540,217

; PRIOR FILING DATE: 2000-03-31

; PRIOR APPLICATION NUMBER: 09/649,167

; PRIOR FILING DATE: 2000-08-23

; NUMBER OF SEQ ID NOS: 60736

; SOFTWARE: Custom

; SEQ ID NO 15295

; LENGTH: 2760

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SIMILAR

; LOCATION: (1282).. (1566)

; OTHER INFORMATION: 73% homologous to Mus musculus Pro-Pol-dUTPase

; OTHER INFORMATION: polyprotein, accession number Y12713, Smith-Waterman Score=381.

US-10-450-763-15295

Query Match 26.3%; Score 691.8; DB 8; Length 2760;

Best Local Similarity 98.3%; Pred. No. 1.1e-155;

Matches 699; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 267 AAGTCCAGCCCTCTCCACAGAGCCGCCAGAGCCCAACAAAGTCTCTCCCTTGACT 326

Db 711 AAGTGAACCGGCACTACAGCAGAGCCGCCAGCCAAACAAAGTCTCTCCCTTGACT 652

Qy 327 TTGAGCAAGAGAGTCTGGGATGTGGAAACCTTTA CCTCTGTCTCTCCCTGGAGCAGCC 386

Db 651 TTGAGCAAGAGAGTCTGGGATGTGGAAACCTTTA CCTCTGTCTCTCCCTGGAGCAGCC 592

Qy 387 CCTCCGAAACCCAGCTCACCAGCGGAAACCAACCGTCCCGAGCAGAACTGCAG 446

Db 591 CCTCCGAAACCCAGCTCACCAGCGGAAACCAACCGTCCCGAGCAGAACTGCAG 532

Qy 447 CAGCTGGAATAGAAATATTCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGAGCAG 506

Db 531 CAGCTGGAATAGAAATATTCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGAGCAG 472

Qy 507 GAGCAAAATCTCTCGTGCCTCAGCATGACAGCCTCTCCCTGTGCGATGACAGTGAAG 566

Db 471 GAGCAAAATCTCTCGTGCCTCAGCATGACAGCCTCTCCCTGTGCGATGACAGTGAAG 412

Qy 567 AACCCATCAAGGCTCTCAAGCTCCGGGACACCTGAGCTCAGGCGTACACTCCNGG 626

Db 411 AACCCATCAAGGCTCTCAAGCTCCGGGACACCTGAGCTCAGGCGTACACTCCNGG 352

Qy 627 AGCAGCCAGGCGGAGGCTCGAAGCTCCAGCTGGAACATGGTGTGCTGGCCCGGGCGGCC 686

Db 351 AGCAGCCAGGCGGAGGCTCGAAGCTCCAGCTGGAACATGGTGTGCTGGCCCGGGCGGCC 292

Qy 687 ACCGACCCCAAGCTATTTCAATGGTGTGAAAGTCCAAATACAGTCTCTTAATGGTGA 746

Db 291 ACCGACCCCAAGCTATTTCAATGGTGTGAAAGTCCAAATACAGTCTCTTAATGGTGA 232

Db 1642 GCTGGAATA----- 1651
Qy 509 GCAATTTCTCGTGCCTCAGCATCGACAGCCTCTCCCTGTCGGATGACAGTGAAGAA 568
Db 1652 -NN 1710
Qy 569 CCCATCAAGGCCTCTCAAGAGTCGCGGACACCCCTGAGCTCAGGCGTACACTCTCTGGAG 628
Db 1711 NNN 1770
Qy 629 CAGCAGCCGAGGCTCGAAGCTCCAGCTGGAACATGTGTGCGCCGCGGCGGCCAC 688
Db 1771 NNN 1830
Qy 689 CGACACCAAGCTATTCAATGGTGTGAAGTCCAAATACAGTCTCTTAATGGTGAACA 748
Db 1831 NNN 1890
Qy 749 CCTGCACATCCGGAGTTCCACCGGTCAAAGTGGGAGACATCGCCACTGGCATCAGCAG 808
Db 1891 NNN 1950
Qy 809 CCAGATCCAGCTGAGCCTTCCAGCTTGGTGCACCAAGAGCGGCGAGCCTGTGCGTAGCA 868
Db 1951 NNN 2010
Qy 869 CATGAGGTCGAGACTCGGGCATCGACTGCAGTGACACTGGGCCCTGATGGCAGCTT 928
Db 2011 NNN 2070
Qy 929 CGCCTGAGCTGAGGGTCAAGCATGGCCAGCTGAGGAGACAGGCCCTTAACCTGCCCCC 988
Db 2071 NNN 2130
Qy 989 ACCGCGGCTCCACACTGCGCGAAAGACAGCCTTCTGCTCGGTGACAGTGTGCGCCTGA 1048
Db 2131 NNN 2190
Qy 1049 AAACACAGGCTCAGCGTTCAGGGGATYTGNCAGCCCCCGCTCARCAGTGGGAA 1108
Db 2191 NNN 2250
Qy 1109 CCAGGCGCTCGCAGCAGCAGTGGGGGCAAGCAGAGATGCTCCAGGATTTAC 1168
Db 2251 NNN 2310
Qy 1169 ANCTTGAGCCCTGCCCCANCCCTCTGADAAAACATYTCGCCACAGTGAAGACAGACA 1228
Db 2311 NNN 2370
Qy 1229 AGGAGATGNCAGGAGTTNNACCTYGGGGAACAAACAGGATCTTTNTTCTGCCCT 1288
Db 2371 NNN 2430
Qy 1289 GCTCCAGTNCGAGTTGGCCTGNACCCTGTTGGANTCAGTGACATTTGTGTCAGANCA 1348
Db 2431 NNN 2490
Qy 1349 GGGAGAGAGCTTCAGCCTGGGTGAGAGGGGTGAGAGGGGTGGCGAGCCCTCGGCCCTCAC 1408
Db 2491 NNN 2550
Qy 1409 NCCAGGCTCTGTGAGAGTGTCAAGTGTGAAGGNCACAAACTCAGGNTTCAGTGCA 1468
Db 2551 NNN 2610
Qy 1469 GAACACAGTNCAGAGGTATGCCCGCCGNTAGGTAAANNGGGGCCCTCTNAACCCCT 1528
Db 2611 NNN 2670
Qy 1529 TGCCTGGCCTNCACTGTCGCCAGCTCACCCTTTTGGGTGAGGGGAAAGATGCT 1588
Db 2671 NNN 2730

Qy 1589 GACCTGGGAAGGCTWCCTGGTAGAATACACACACTTTTCAGGTTGTTGCAACACAGG 1648
Db 2731 NNN 2790
Qy 1649 TCCTGAGTTGACCTCTGGTTCAGCAAGGACAAAGAGGTGTGTAAGTGAAGTGTCT 1708
Db 2791 NNN 2850
Qy 1709 CAGTNCACAGCATGTGCCCTTTGCTGTGGCTACCACTCTTCCCAGAGCAGCAGC 1768
Db 2851 NNN 2910
Qy 1769 CCCGAGCCCTTCAGGCCACGACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTG 1828
Db 2911 NNN 2970
Qy 1829 TAAAGGTGAAGGTGATGAGGATATGCTGCACAGGAACAGTCTGTGGATGGACATGATC 1888
Db 2971 NNN 3030
Qy 1889 AGTCTNAAGGNAAGCAGCAGAGAGAGCGYTCGGCGCCCGAGGCCACTNATCAGT 1948
Db 3031 NNN 3084
Qy 1949 GTNCCAGCGTCTGCTGTTNCCCAGNAGCAGCTNAGNCACTCANCACTGACACTNCA 2008
Db 3085 CAGCTCCAGCGTGGGGCCCCCAGAGCAGCT-----CACAAACACTCA 3133
Qy 2009 CTTNGCCCTGCTGCTGCGCCANGAGGGTATCTGCGNAGCGGACTTTGCACTCTGATGNA 2068
Db 3134 CCTGCTGCTGCCCC- -GGGCGGAGGGTACTGCTG-ACAGCACTTTGAC- TCTGATG-A 3188
Qy 2069 CCTCAAGCACTTTTCAGGCTGCTGCTGNNGGCAGGNCAGGNCAGGNCAGTGCAN 2128
Db 3189 CCTCAAGCACTTTTCAGGCTGCTGCTGNNGGCAGGNCAGGNCAGGNCAGTGCAN 2188
Qy 2129 CTGTAGGAGCATANGCAANGCCAGGAGATGGGTGNAAGGGANCAAGTCTTTCAGCTGT 2188
Db 3234 ACTCTGTAGGAGCATAGCAGTCCAGGAGATGGGTGAGGG- -ACAGTCTTTCAGCTGT 3289
Qy 2189 CCANCATGATGCTGACTNCTCAAACTCTTNNCAGNATTTCTCTAAGAATAGCAGCCC 2248
Db 3290 CTA- CATGATGCTGACTCTCAAACTCTCT- -CCAGATTTCTCTAATAGCAACCCAC- - 3342
Qy 2249 CTTNCCCATTTGCCCACTTTCAGCTTCTTCCAGGGGAGCTANCTCAGGACTCAGCT 2308
Db 3343 ---TTCCCATTTGCCCACTTTCAGCTTCTTCCAGGGGAGCTA-CTCAGGACTCATGT 3398
Qy 2309 AGCATTAATCAGCTGTGNAATCGTTCAGGGGTGTGCTGAGCCTCAACCTCTGGGGCA 2368
Db 3399 AGCATTAATCAGCTGTGNAATCGTTCAGGGGTGTGCTGAGCCTCAACCTCTGGGGCG 3458
Qy 2369 GGGGAGCGGAGACTCCGTGGGAGAGCTCATTTCCACATCTTCCAGAGACAGCCTTTNG 2428
Db 3459 GGGGAGCGGAGACTCTTGGGAGAGCTCATCCCCAGCTCTG- -NNNNAAGCTTTG 3516
Qy 2429 TCCAGCTCTCACATTTGAGTCAAGTCTGCTCCGGGAGAGAGCCCGGCCCGCCAGCAT 2488
Db 3517 TCCAGCTCTCACAGCGAGCAGACTGCTCCGGGAGAGAGCCCGGCCCGCCAGCAT 3576
Qy 2489 AAAGAACTGCAAGCTTGTACTGCAAGTCTGGGTGTGAGAACTCTTTGTGAAGCAATA 2548
Db 3577 AAAGAACTGCAAGCTTGTACTGCAAGTCTGGGTGTGAGAACTCTTTGTGAAGCAATA 3634
Qy 2549 AAGTTGGGTGATGACAGATGTT 2572
Db 3635 AAGTTGGGTGATGACAGATGTT 3658

RESULT 12
US-60-680-473-32424
; Sequence 32424, Application US/60680473

GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary
; TITLE OF INVENTION: Fragments of, Cynomolgus Monkey Genes and the Use Thereof
; FILE REFERENCE: 21590290000
; CURRENT APPLICATION NUMBER: US/60/680,473
; NUMBER OF SEQ ID NOS: 48714
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 32424
; LENGTH: 3658
; TYPE: DNA
; ORGANISM: Macaca Mulatta
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(3658)
; OTHER INFORMATION: n = A,T,C or G
US-60-680-473-32424

Query Match 11.2%; Score 295.6; DB 15; Length 3658;
Best Local Similarity 28.7%; Pred. No. 1.4e-60;
Matches 645; Conservative 1; Mismatches 1491; Indels 107; Gaps 15;
QY 329 GAGCAAGGAGGAGTCTGGGATGTGGGAACCTTACCTCTGTCTCTCCCTGGAGCCAGCCCC 388
DB 1522 GAGCAAGGAGGAGTCTGGGATGTGGGAACCTTACCTCTGTCTCTCCCTGGAGCCAGCCCC 1581
QY 389 TGCCAGAAACCCAGCTCACAGAGCGGAAGCAACCTCCCGAGCAGGAAGTGCAGCA 448
DB 1582 TGCCAGAAACCCAGCTCACAGAGCGGAAGCAACCTCCCGAGCAGGAAGTGCAGCA 1641
QY 449 GCTGGAATAGAAATATTCCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGCAGGA 508
DB 1642 GCTGGAATA----- 1651
QY 509 GCAAAATCTCTGTGCTCAGCATCGACAGCCTCTCCCTGCGGATGACAGTGAGAAGAA 568
DB 1652 -NN 1710
QY 569 CCATCAAGAGCCTCTCAAAGCTCGCGGACACCTCAGCTCAGGCGCTACACTCTGGAG 628
DB 1711 NNN 1770
QY 629 CAGCCAGCCGAGGCTCGAAGCTCCAGCTGGAACATGGTGTGCGCCCGGGCGGCCAC 688
DB 1771 NNN 1830
QY 689 CGACACCCCAAGTATTCTCAATGGTGTGAAGTCAAAATACAGTCTCTTAATGGTGAACA 748
DB 1831 NNN 1890
QY 749 CTTGCAATCCGGAGGTTCCACCGGTCAAAGTGGGAGACATCGCCACTGGCATCAGCAG 808
DB 1891 NNN 1950
QY 809 CCAGATCCAGCTGCGACCTTACGCTTGGTCAACCAAGACGGGCGAGCTGTTCGCTACGA 868
DB 1951 NNN 2010
QY 869 CATGGAGGTGCCAGATCGGGCATCGACCTGCACTGCACTGGCCCCCTGATGGCAGCTT 928
DB 2011 NNN 2070
QY 929 CGCTGGAGCTGGAGGTCAAGCATGGCCAGCTGGAGAAACAGGCCCTTAACCTGCCCTCC 988
DB 2071 NNN 2130

QY 989 ACCGCCGGCTCCACACTGCCGAAAGCAGCGCTTCTCTGCTCGGTGCACGATGCTGCCCTGA 1048
DB 2131 NNN 2190
QY 1049 AAACACAGGCTCAGCGCTTCCAGGGGATGTGNCACAGCCCCCGGCTCARCAGTGGAA 1108
DB 2191 NNN 2250
QY 1109 CCAGGGCTCCGACAGCAGCAGGAGTGGGGCAAGCAGAGAAATGCTCCAGGATTTTCAC 1168
DB 2251 NNN 2310
QY 1169 ANCTGAGCCCTGCCCCCANCCTCTGTGAADAAACAATTCGCCCACTGTAAGAGACAGA 1228
DB 2311 NNN 2370
QY 1229 AGGAGGATGNCAGGAGGATTNACCTYGGGGAACAAACAGGGATCTTTTCTGCCCCCT 1288
DB 2371 NNN 2430
QY 1289 GCTCCAGTNCAGTTCGCCCTGNACCCGCTTGGANTCAGTGACCATTTGTTGGCAGACAG 1348
DB 2431 NNN 2490
QY 1349 GGGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTCGCGCCCTCACCCCT 1408
DB 2491 NNN 2550
QY 1409 NCAGGCTGCTGTGNAGAGTGTCAAGTGTGAAGGCCCAAAANCTCAGGTTTCAGTGCA 1468
DB 2551 NNN 2610
QY 1469 GAACAGGTCACAGGATGCGCCCGCCGNTAGTTAAANGGGGGCCCTCTNAAACCCCT 1528
DB 2611 NNN 2670
QY 1529 TGCTTGGCCCTNACCTTNGGCCAGCTCANCCTCTTTGGGTGTAGGGGAAAGAAATGCCT 1588
DB 2671 NNN 2730
QY 1589 GACCTGGGAAGGTGCTGCTGGTAGAATAACACACATTTTCAGGTTGTGCAACACAGG 1648
DB 2731 NNN 2790
QY 1649 TCCTGAGTTGACCTCTGGTTTCAGCAAGGACCAAGAGGTGTCTAAGTGAAGTGTCT 1708
DB 2791 NNN 2850
QY 1709 CAGTNCCTCCAGACATGTGCCCCCTTTGCTGTGCTACCACTCTTCCCGCAGAGCAGGC 1768
DB 2851 NNN 2910
QY 1769 CCCGAGCCCTTCAGGCCAGCACTGCCCCAGACTCGCTGGCACTGCTCCCTCATCTG 1828
DB 2911 NNN 2970
QY 1829 TAAAGTGAAGGTGATGCAGGATATGCTGTGACAGGAACAGTCTGTGTGATGGAATGC 1888
DB 2971 NNN 3030
QY 1889 AGTGTNAAAGNAAGCAGCAGAGAGAGAGCTCCGCGGCCCGCCAGNCCCACTNATCAGT 1948
DB 3031 NNN 3084
QY 1949 GTNCCAGCTGCTGTTGTTTCCCGAGAGCACTGTCAGNCACTCANCACTGACACTNAC 2008
DB 3085 CAGCGTCCAGCTGCGGGCCCCCAGAGCAGCT-----CACAAACACTCA 3133
QY 2009 CTTNGCCCTGCTGGCCAGAGGCTACTGCGNACGCACTTTCGCACTTGTGATGNA 2068
DB 3134 CCCTGCCCTGCCCT--GGGCGGAGGCTACTGCTG-ACAGCACTTTGAC-TCTGATG-A 3188

Qy	2069	CCTCAAGCATTTCATGGCTTGGCTTCTNNGGCAGGNCAGGNCAGGNCAGTGCAN	2128
Db	3189	CCTCAAGCATTTCATGGCTTGGCTTCTNNGGCAGGNCAGGNCAGGNCAGTGCAN	3233
Qy	2129	CTGTAGGAGCATANGCAAGCCAGGAGATGGGGTGNAAAGGANCACAGTCTTTCAGCTGT	2188
Db	3234	ACTCTGTAGGAGCATANGCAAGCCAGGAGATGGGGTGNAAAGGANCACAGTCTTTCAGCTGT	3289
Qy	2189	CCAMCATCATGTGATCTNCCCTCAAACTCTTNNCCAGNATTTCTTAAGAATACGACCC	2248
Db	3290	CTA-CATGATGTGACTCTCAAACTCTTNNCCAGNATTTCTTAAGAATACGACCC	3342
Qy	2249	CCTNCCCATTTGCCAGCTAGCTCTTCTCCAGGGAGTANCTCAGGACTCAGT	2308
Db	3343	---TTCCTCATTTGCCAGCTAGCTCTTCTCCAGGGAGTANCTCAGGACTCAGT	3398
Qy	2309	AGCATTAAATCAGCTGTGNAATCGTCCAGGGGGTGTCTGCTAGCTCAACCTCTCTGGGGCA	2368
Db	3399	AGCATTAAATCAGCTGTGNAATCGTCCAGGGGGTGTCTGCTAGCTCAACCTCTCTGGGGCA	3458
Qy	2369	GGGGACGCGAGACTCCGTGGAGAGTCAATTTCCCAATCTTTGCCAAGAAGCTTTTG	2428
Db	3459	GGGGACGCGAGACTCCGTGGAGAGTCAATTTCCCAATCTTTGCCAAGAAGCTTTTG	3516
Qy	2429	TCCAGCTGTCCATTTGAGTCAAGCTGCTCCGGGGAGAGCCCGCCCGCCAGCACAT	2488
Db	3517	TCCAGCTGTCCATTTGAGTCAAGCTGCTCCGGGGAGAGCCCGCCCGCCAGCACAT	3576
Qy	2489	AAAGAACTGCAGCTTGGTACTGTCAGAGTCTGGGTGTAGAGAACTTTTGTAAAGCAATA	2548
Db	3577	AAAGAACTGCAGCTTGGTACTGTCAGAGTCTGGGTGTAGAGAACTTTTGTAAAGCAATA	3634
Qy	2549	AAAGTTGGGGTGCATGACAAATGTT	2572
Db	3635	AAAGTTGGGGTGCATGACAAATGTT	3658
RESULT 13			
US-60-544-10785/c			
; Sequence 10785, Application US/60680544			
; GENERAL INFORMATION:			
; APPLICANT: Cooper, Matthew			
; APPLICANT: Kinch, Deborah			
; APPLICANT: Rosenberg, Michael			
; APPLICANT: Subramaniam, S. Sai			
; APPLICANT: Szak, Suzanne			
; APPLICANT: Li, Huo			
; APPLICANT: Bandaru, Raj			
; APPLICANT: Derbel, Maher			
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary			
; FILE REFERENCE: 2159029000			
; CURRENT FILING DATE: 2005-05-13			
; NUMBER OF SEQ ID NOS: 48714			
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0			
; SEQ ID NO 10785			
; LENGTH: 761			
; TYPE: DNA			
; ORGANISM: Macaca Fascicularis			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (1)...(761)			
; OTHER INFORMATION: n = A, T, C or G			
US-60-544-10785			
Query Match 9.8%; Score 257.4; DB 15; Length 761;			
Best Local Similarity 59.8%; Pred. No. 1.4e-51;			
Matches 490; Conservative 1; Mismatches 269; Indels 59; Gaps 11;			
Qy	1719	GACATGTGCCCCCTTTGCTGTGCTACCACTCTTCCCGAGAGCAGGCCCCGAGCCCC	1778
Db	761	GACACATGNNNNNGCTGTGCTATCANNNNTCCCGAGCAGCAGNNNNNGAGCCCC	702

Qy	1779	TTCCAGCCCGCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCTCATCTCTAAAGGTGAA	1838
Db	701	TTCCAGCCCGCAG--NN	644
Qy	1839	GGGTGATGCAAGTATGCTGACAGAAACAGTCTGTGGATGGAATGATCATGCTGCTNAAG	1898
Db	643	NN	586
Qy	1899	GNAAAGCAGCAGAGAGAGCGTCCGGCGCCCGCCAGNCCCACCTNATCAGGTGTCACGCGT	1958
Db	585	GGAAAGCAGCAGAGAGAGCGTCCGGCGCCCGCCAGNCCCACCTNATCAGGTGTCACGCGT	532
Qy	1959	GCTNGGTTCNCCAGNAGCAGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	2018
Db	531	CGTCCGGGCCCCCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	488
Qy	2019	CCCTNCGCCAGAGGGGTCTGCGGACAGGCACTTTTGCACTGTCAGTGTGAGTGTGAGTGTGAG	2078
Db	487	CCCTNCGCCAGAGGGGTCTGCGGACAGGCACTTTTGCACTGTCAGTGTGAGTGTGAGTGTGAG	428
Qy	2079	CTTTTCATGGCTGCTCTTNNCGCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAG	2138
Db	427	CTTTTCATGGCTGCTCTTNNCGCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAG	380
Qy	2139	CATANGCAAGCCAGGAGATGGGGTGNAAAGGAGCAGCTCTTGAGCTGTCCANCACTGCA	2198
Db	379	-----ANN	328
Qy	2199	TGTGACTNCTCAAACTCTTNNCGCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAG	2258
Db	327	TGTGACTNCTCAAACTCTTNNCGCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAG	277
Qy	2259	TTGCCCGCAGCTTACGCTCTTCCCGAGGAGTANCTCAGGACTCAGTGTGAGTGTGAGTGTGAGTGTGAG	2318
Db	276	TTGCCCGCAGCTTACGCTCTTCCCGAGGAGTANCTCAGGACTCAGTGTGAGTGTGAGTGTGAGTGTGAG	218
Qy	2319	CAGCTGTGNAATCGTCAGGGGGTGTCTGCTAGCTCAACCTCTCTGGGCGAGGGGAGCGCG	2378
Db	217	CAGCTGTGNAATCGTCAGGGGGTGTCTGCTAGCTCAACCTCTCTGGGCGAGGGGAGCGCG	158
Qy	2379	AGACTCCCGTGGGAGAGCTCATTTCCACATCTTCCAGAGCAGGCTTTTNGTCCAGCTGTC	2438
Db	157	AGACTCCCGTGGGAGAGCTCATTTCCACATCTTCCAGAGCAGGCTTTTNGTCCAGCTGTC	99
Qy	2439	CACATTTGAGTCAGACTGCTCCCGGAGAGAGCCCGCCCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	2498
Db	98	CACATTTGAGTCAGACTGCTCCCGGAGAGAGCCCGCCCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	39
Qy	2499	AGCCTTGGTACTGAGAGTCTGGGTGTGAGAGAACTCTT	2537
Db	38	AGCCTTGGTACTGAG--AGTGTGGGTGTGAGAGAACTCTT	2
RESULT 14			
US-60-544-11056			
; Sequence 11056, Application US/60680544			
; GENERAL INFORMATION:			
; APPLICANT: Cooper, Matthew			
; APPLICANT: Kinch, Deborah			
; APPLICANT: Rosenberg, Michael			
; APPLICANT: Subramaniam, S. Sai			
; APPLICANT: Szak, Suzanne			
; APPLICANT: Li, Huo			
; APPLICANT: Bandaru, Raj			
; APPLICANT: Derbel, Maher			
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary			
; FILE REFERENCE: 2159029000			
; CURRENT FILING DATE: 2005-05-13			
; NUMBER OF SEQ ID NOS: 48714			
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0			
; SEQ ID NO 11056			
; LENGTH: 761			
; TYPE: DNA			
; ORGANISM: Macaca Fascicularis			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (1)...(761)			
; OTHER INFORMATION: n = A, T, C or G			
US-60-544-10785			
Query Match 9.8%; Score 257.4; DB 15; Length 761;			
Best Local Similarity 59.8%; Pred. No. 1.4e-51;			
Matches 490; Conservative 1; Mismatches 269; Indels 59; Gaps 11;			
Qy	1719	GACATGTGCCCCCTTTGCTGTGCTACCACTCTTCCCGAGAGCAGGCCCCGAGCCCC	1778
Db	761	GACACATGNNNNNGCTGTGCTATCANNNNTCCCGAGCAGCAGNNNNNGAGCCCC	702

; SEQ ID NO 11056
; LENGTH: 761
; TYPE: DNA
; ORGANISM: Macaca Fascicularis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(761)
; OTHER INFORMATION: n = A,T,C or G
US-60-680-544-11056

Query Match 9.8%; Score 257.4; DB 15; Length 761;
Best Local Similarity 59.8%; Pred. No. 1.4e-51;
Matches 490; Conservative 1; Mismatches 269; Indels 59; Gaps 11;

Qy 1719 GACATGTGCCCTTTGCTGCTGTACCACTCTTCCCGCAGAGCAGCGCCGAGCCCC 1778
Db 1 GACACATGNNNNNGCTGCTGGCTATCANNNTCCCGCAGCGCAGGNNNNNGAGCCCC 60

Qy 1779 TTCAGGCCAGACTGCCCGAGCTCGCTGGCACTAGTTCCTCATCTCTAAAGGTGAA 1838
Db 61 TTCAGGCCAG--NN 118

Qy 1839 GGGTGATGCAGGATATCCCTGACAGGAACAGTCTGTGGATGGACATGATCAGTCTNAAG 1898
Db 119 NNNNNNNNNNNNNCCCTGACAGGAACGGTCTGTAGATGGACATGATCAATACTAA-- 176

Qy 1899 GNAAGCAGCAGAGAGAGCGYTCCGGCGCCCGCAGNCCCACTNATCAGTGTNCCAGCGT 1958
Db 177 GGAAGCAGCAGAGAGAGCGTCCGGCACCCCGAGCCCC--ACTGTGAGGTCCAG 230

Qy 1959 GCTNGGTTNCCCGCAGNAGCAGCTNCAAGNATCANCATGACACTNCACTTNGCCCTG 2018
Db 231 CGTGGCGGCCCGCAGAGCAGCTCACGA-----ACACTCACCTG 274

Qy 2019 CCCTNGGCCANGAGGATCTGCCGACGCACTTTGCACNTCTGATGNACCTCAAAGCA 2078
Db 275 CCGCGCCCGCGGAGGGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 334

Qy 2079 CTTTCATGGCTNGCCCTCTNNGCAGGNCAGGNCAGGNCAGTGCANCTGTAGNAG 2138
Db 335 CTTTCATGGCTG---CCCTCTGGCAGGCGAGTGCAGTCTGTAGGAGC----- 382

Qy 2139 CATANGAANGCCAGGAGATGGGGTGNAAAGGANCACAGTCTTGAGCTGTCCANCAATGCA 2198
Db 383 -----ANNNGCA 434

Qy 2199 TGTGACTNCTCAAACCTCTTNNCCAGNATTTCTTAAGATAGCANCCTTNCCTTNCCTCA 2258
Db 435 TGTGACTCTCAAACCTCT---CCAGATTTCTTAATAGCAACCACT---TTCCCTCA 485

Qy 2259 TTGCCCCAGCTTAGCCTCTTCTCCAGGGAGCTTANCTCAGGACTCAGTAGCATTAAT 2318
Db 486 TTGCCCCAGCTTAGCCTCTTCTCCAGGGAGCTA-CTCAGGACTCATGTAGCATTAAGT 544

Qy 2319 CAGCTGTGNAATCGTCAGGGGTGTCTGCTAGCCTCAACCTCTGGGGCAGGGGAGCGCG 2378
Db 545 CAGCTGTGAAATTGTGAGGGGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 604

Qy 2379 AGACTCGTGGGAGAGCTCATTTCCACATCTTGGCAGAGCAGCTTTTNGTCCAGCTGTC 2438
Db 605 AGACTCTTGGGAGAGCTCATTTCCACGCTTTGGCAAGACAGCTTTT-GTCCAGCTGTC 663

Qy 2439 CACATTGAGTCAGACTCTCCGGGAGAGAGCGCGCCCGCCCGCAGACATAAAGAACTGC 2498
Db 664 CACAGGAGCAGACTGCTCCGGGAGAGAGCGCGCCCGCCCGCAGACATAAAGAACTGC 723

Qy 2499 AGCCTGGTACTGAGAGTCTGGGTGTGAGAGAACTCTT 2537
Db 724 AGCCTGGTACTGC--AGTGTGGGTGTGAGAGAACTCTT 760

RESULT 15

US-60-680-473-10785/c

; Sequence 10785, Application US/60680473
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary to
; FILE REFERENCE: 21590290000
; CURRENT APPLICATION NUMBER: US/60/680,473
; CURRENT FILING DATE: 2005-05-13
; NUMBER OF SEQ ID NOS: 48714
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 10785
; LENGTH: 761
; TYPE: DNA
; ORGANISM: Macaca Fascicularis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(761)
; OTHER INFORMATION: n = A,T,C or G
US-60-680-473-10785

Query Match 9.8%; Score 257.4; DB 15; Length 761;
Best Local Similarity 59.8%; Pred. No. 1.4e-51;
Matches 490; Conservative 1; Mismatches 269; Indels 59; Gaps 11;

Qy 1719 GACATGTGCCCTTTGCTGCTGTACCACTCTTCCCGCAGAGCAGCGCCGAGCCCC 1778
Db 761 GACACATGNNNNNGCTGCTGGCTATCANNNTCCCGCAGCGCAGGNNNNNGAGCCCC 702

Qy 1779 TTCAGGCCAGACTGCCCGAGCTCGCTGGCACTAGTTCCTCATCTCTAAAGGTGAA 1838
Db 701 TTCAGGCCAG--NN 644

Qy 1839 GGGTGATGCAGGATATCCCTGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAG 1898
Db 643 NNNNNNNNNNNNNCCCTGACAGGAACGGTCTGTAGATGGACATGATCAATACTAA-- 586

Qy 1899 GNAAGCAGCAGAGAGAGCGYTCCGGCGCCCGCAGNCCCACTNATCAGTGTNCCAGCGT 1958
Db 585 GGAAGCAGCAGAGAGAGCGTCCGGCACCCCGAGCCCC--ACTGTGAGGTCCAG 532

Qy 1959 GCTNGGTTNCCCGCAGNAGCAGCTNCAAGNATCANCATGACACTNCACTTNGCCCTG 2018
Db 531 CGTGGCGGCCCGCAGAGCAGCTCACGA-----ACACTCACCTG 488

Qy 2019 CCCTNGGCCANGAGGATCTGCCGACGCACTTTGCACNTCTGATGNACCTCAAAGCA 2078
Db 487 CCGCGCCCGCGGCGGAGGGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 428

Qy 2079 CTTTCATGGCTNGCCCTCTNNGCAGGNCAGGNCAGGNCAGTGCANCTGTAGNAG 2138
Db 427 CTTTCATGGCTG---CCCTCTGGCAGGGCAGGTGAGTGTCTGTAGGAGC----- 380

Qy 2139 CATANGAANGCCAGGAGATGGGGTGNAAAGGANCACAGTCTTTGAGCTGTCCANCAATGCA 2198
Db 379 -----ANNNGCA 328

Qy 2199 TGTGACTNCTCAAACCTCTTNNCCAGNATTTCTTAAGATAGCANCCTTNCCTTNCCTCA 2258
Db 327 TGTGACTCTCAAACCTCTCT---CCAGATTTCTTAATAGCAACCACT---TTCCCTCA 277

Qy 2259 TTGCCCCAGCTTAGCCTCTTCTCCAGGGAGCTTANCTCAGGACTCAGCTAGCATTAAT 2318
Db 276 TTGCCCCAGCTTAGCCTCTTCTCCAGGGAGCTA-CTCAGGACTCATGTAGCATTAAGT 218

Qy 2319 CAGCTGTGNAATCGTTCAGGGGTGTCTGCTAGCCTCAAACCTCTTGGGCGAGGGAGCGCG 2378
Db 2319 CAGCTGTGNAATCGTTCAGGGGTGTCTGCTAGCCTCAAACCTCTTGGGCGAGGGAGCGCG 2378
Db 545 CAGCTGTGAAATTGTGAGGGGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 604
Db 2379 AGACTCGTGGGAGAGCTCATTTCCACATCTTGGCAGAGCAGCTTTTNGTCCAGCTGTC 2438
Db 605 AGACTCTTGGGAGAGCTCATTTCCACGCTTTGGCAAGACAGCTTTT-GTCCAGCTGTC 663
Db 2439 CACATTGAGTCAGACTCTCCGGGAGAGAGCGCGCCCGCCCGCAGACATAAAGAACTGC 2498
Db 664 CACAGGAGCAGACTGCTCCGGGAGAGAGCGCGCCCGCCCGCAGACATAAAGAACTGC 723
Qy 2499 AGCCTGGTACTGAGAGTCTGGGTGTGAGAGAACTCTT 2537
Db 724 AGCCTGGTACTGC--AGTGTGGGTGTGAGAGAACTCTT 760

Db 217 CAGCTGTGAATTGTCAAGGGGGTGTCAAGCTAGCCTCAACCTCTCTGGGGCGGGGACGCCG 158
QY 2379 AGACTCCGTGGAGAGCTCAATCCACATCTTGGCCAAGACAGCCCTTTNGTCCAGCTGTC 2438
Db 157 AGACTCCTTTGGAGAGCTCATCCCACTGCTTGGCCAAGACAGCCCTT-GTCCAGCTGTC 99
QY 2439 CACATTGAGTCAGACTGTCTCCGGGGAGAGAGCCCGGCCCCAGCACATAAAGAACTGC 2498
Db 98 CACAGCGAGCCAGACTGTCTCCGGGGAGACAGCCCGGCCCCAGCACATAAAGAACTGC 39
QY 2499 AGCCTTGGTACTGACAGACTCTGGGTGTAGAGAACTCTT 2537
Db 38 AGCGTTGGTACTGC--AGTGTGGTGTAGAGAACTCTT 2

Search completed: June 11, 2005, 05:07:11
Job time : 1021.45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:18:24 ; Search time 3562.59 Seconds
(without alignments)
17042.190 Million cell updates/sec

Title: US-09-155-676B-4
Perfect score: 1253
Sequence: 1 cattggaggtcacgcggtggc.....ggggggcggtaccacatttt 1253

Scoring table: IDENTITY NUC
Gapop 10'0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.on.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1252	99.9	1253	6	A66648	A66648 Sequence 4
2	1021.8	81.5	1202	9	AB094093	AB094093 Homo sapi
3	723.8	57.8	1175	10	BC055929	BC055929 Mus muscu
4	489	39.0	104913	9	H3718P11	AL109983 Human DNA
5	489	39.0	192748	2	AC011089	AC011089 Homo sapi
6	444.4	35.5	500	6	CQ700802	CQ700802 Sequence
7	400	31.9	601	11	BV167722	BV167722 sqnm6310
8	400	31.9	1191	11	BV177918	BV177918 sqnm97502
9	392	31.3	401	6	AX071560	AX071560 Sequence
10	327.4	26.1	486	6	CQ752826	CQ752826 Sequence
11	309	24.7	1390	5	CR406372	CR406372 Gallus ga
12	297	23.7	421	6	BD058555	BD058555 Secreted
13	281.4	22.5	433	6	CQ727443	CQ727443 Sequence
14	263	21.0	263	6	CQ728296	CQ728296 Sequence
15	244.8	19.5	374	6	AX906736	AX906736 Sequence
16	244.8	19.5	374	6	BD042269	BD042269 Sequence
17	186.4	14.9	254	6	CQ707871	CQ707871 Sequence
18	175.8	14.0	158634	2	AF000909	AF000909 Homo sapi
19	175.8	14.0	165394	2	AP000877	AP000877 Homo sapi

20	175.8	14.0	178341	9	AP003392	AP003392 Homo sapi
21	175.8	14.0	196511	2	AP000854	AP000854 Homo sapi
c 22	175.8	14.0	244847	2	AP002357	AP002357 Homo sapi
c 23	173	13.8	165394	2	AP000877	AP000877 Homo sapi
c 24	173	13.8	196511	2	AP000854	AP000854 Homo sapi
c 25	173	13.8	244847	2	AP002357	AP002357 Homo sapi
c 26	160.8	12.8	3793	9	AB061844	AB061844 Homo sapi
c 27	134.6	10.7	246370	2	AC107575	AC107575 Rattus no
c 28	134.6	10.7	248500	2	AC104400	AC104400 Rattus no
c 29	134.6	10.7	258426	2	AC105645	AC105645 Rattus no
30	132.6	10.6	185902	10	AC122428	AC122428 Mus muscu
31	122.8	9.8	328	6	AX918334	AX918334 Sequence
32	122.8	9.8	328	6	BD053867	BD053867 Sequence
33	103.4	8.3	1727	5	BX933109	BX933109 Gallus ga
34	102.4	8.2	158634	2	AP000909	AP000909 Homo sapi
35	83.4	6.7	97	6	AX899765	AX899765 Sequence
36	83.4	6.7	97	6	BD035298	BD035298 Sequence
c 37	67.6	5.4	701	10	BC079541	BC079541 Mus muscu
c 38	61.6	4.9	1906	6	A66645	A66645 Sequence 1
c 39	60.4	4.8	7218	6	I66494	I66494 Sequence 14
c 40	55.2	4.4	125020	9	AF429315	AF429315 Homo sapi
c 41	54.6	4.4	2000	6	AX655393	AX655393 Sequence
c 42	53.2	4.2	2631	6	A66647	A66647 Sequence 3
c 43	50.8	4.1	198378	2	AC108256	AC108256 Rattus no
c 44	49	3.9	97839	2	AC141520	AC141520 Rattus no
45	47.8	3.8	1080	9	AF035606	AF035606 Homo sapi

ALIGNMENTS

RESULT_1
LOCUS A66648 1253 bp DNA linear PAT 29-MAR-1999
DEFINITION Sequence 4 from Patent WO9737016.
ACCESSION A66648
VERSION A66648.1 GI:4538140
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1253)
AUTHORS Wallach,D., Malinin,N., Boldin,M., Kovalenko,A. and Mett,I.
TITLE MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR PREPARATION AND USE
JOURNAL Patent: WO 9737016-A 4 09-OCT-1997;
YEDA RES & DEV (IL)
COMMENT Other publication AU 2175597 19971022.
FEATURES
source Location/Qualifiers
1..1253
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN	Query Match 99.9%; Score 1252; DB 6; Length 1253; Best Local Similarity 100.0%; Pred. No. 0; Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CATTGGAGTCAACGCGGTGGCGGCCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60
Db	1 CATTGGAGTCAACGCGGTGGCGGCCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60
Qy	61 GATTTCGAGCCACGAGGCGCCCTTCTTGTGTCGCGGCACGTTTACAGCCGCAAGCAC 120
Db	61 GATTTCGAGCCACGAGGCGCCCTTCTTGTGTCGCGGCACGTTTACAGCCGCAAGCAC 120
Qy	121 CCACGGCAGCTGAAGGAGGCTTTTGAGAGCTCTCTCCCGAGGTGAGGGCGGCCGCA 180
Db	121 CCACGGCAGCTGAAGGAGGCTTTTGAGAGCTCTCTCCCGAGGTGAGGGCGGCCGCA 180
Qy	181 GGCCATCCGCGCGCTCAGGTGGAGCGCTATGTGCCCGAAACACGAGCGATCTCTGTGTG 240

Db 181 GGCCATCCGCGCGCTCAGGTGGAGCGCTATGTGCCGAAACACGAGCGATGCTGCTGGTG 240
Qy 241 CCTGTGCTGCGGCTGTGAGTCCGGGAACACCTGAGCCATGCAACCTGAGCGTCTGTA 300
Db 241 CCTGTGCTGCGGCTGTGAGTCCGGGAACACCTGAGCCATGCAACCTGAGCGTCTGTA 300
Qy 301 CGGGGGGCTGTGGAGCTCTGGCCAGCCAGAGCACCAAGAACCAACCAAAATTCG 360
Db 301 CGGGGGGCTGTGGAGCTCTGGCCAGCCAGAGCACCAAGAACCAACCAAAATTCG 360
Qy 361 GTGGAGAACAAAGCTGAGGTCCAGATGAAGAGAGTTCCTGTCATCTCCCGAGATTA 420
Db 361 GTGGAGAACAAAGCTGAGGTCCAGATGAAGAGAGTTCCTGTCATCTCCCGAGATTA 420
Qy 421 TCGCGATTCAAGAAATCATGTCGAAGTTTGGATTCTATGAAGAAAGGAGATTA 480
Db 421 TCGCGATTCAAGAAATCATGTCGAAGTTTGGATTCTATGAAGAAAGGAGATTA 480
Qy 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGAGTGGT 540
Db 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGAGTGGT 540
Qy 541 TCGGCTGTCTTAGAGCTCAGGCAAGTGCAGACCCAGAGAGGCTCTTCAGACCTAG 600
Db 541 TCGGCTGTCTTAGAGCTCAGGCAAGTGCAGACCCAGAGAGGCTCTTCAGACCTAG 600
Qy 601 AAGCTGGAAGGATGAAAGAGTGTGTTGTTTCCAGCTTACAGAGCCCTCAAAATTGA 660
Db 601 AAGCTGGAAGGATGAAAGAGTGTGTTTCCAGCTTACAGAGCCCTCAAAATTGA 660
Qy 661 CTGCGCACCAGCTCCAGAGCTTGAAGTGGATGGAGCAGGACCATCTCTGACATTCATGG 720
Db 661 CTGCGCACCAGCTCCAGAGCTTGAAGTGGATGGAGCAGGACCATCTCTGACATTCATGG 720
Qy 721 CCATCAGGATATACAGGAGTGTGTAACATCCATCAGGTGCCACCTCCCTGGATGAT 780
Db 721 CCATCAGGATATACAGGAGTGTGTAACATCCATCAGGTGCCACCTCCCTGGATGAT 780
Qy 781 CCAAGTGAAGATACATTTGCTGGGAACCAAGAAATAGACCATCTATGAAGATTTCT 840
Db 781 CCAAGTGAAGATACATTTGCTGGGAACCAAGAAATAGACCATCTATGAAGATTTCT 840
Qy 841 TAAAGAAAAGGAAAACAGAGTTGAAAAAACTCCCCCAGACCCAGAGTTGGGCCCAACTT 900
Db 841 TAAAGAAAAGGAAAACAGAGTTGAAAAAACTCCCCCAGACCCAGAGTTGGGCCCAACTT 900
Qy 901 TGATCAGCTCAGGACAGTGCAGGTGGCTGGCTCTTTTGGGCGCGCTCTGGAATA 960
Db 901 TGATCAGCTCAGGACAGTGCAGGTGGCTGGCTCTTTTGGGCGCGCTCTGGAATA 960
Qy 961 ATGGAGCGCGTGGAGTCCAGACATCACTCCAAACTGAAGCTGCAGCAATGAAGAG 1020
Db 961 ATGGAGCGCGTGGAGTCCAGACATCACTCCAAACTGAAGCTGCAGCAATGAAGAG 1020
Qy 1021 CAGTCAATACAGAAAAAGCTTAATCATGCTCTCTACCAACTACCATGAGGCTAAAGGC 1080
Db 1021 CAGTCAATACAGAAAAAGCTTAATCATGCTCTCTACCAACTACCATGAGGCTAAAGGC 1080
Qy 1081 AAAGTCAACAAACCCCTATTATACCTTCAACCCAAATTTCTTTATCATTTCTTTAG 1140
Db 1081 AAAGTCAACAAACCCCTATTATACCTTCAACCCAAATTTCTTTATCATTTCTTTAG 1140
Qy 1141 GAAACAGATACATCAATTTGATTATTAATGAATTTTATTTTCGGCTTCGTCGCT 1200
Db 1141 GAAACAGATACATCAATTTGATTATTAATGAATTTTATTTTCGGCTTCGTCGCT 1200
Qy 1201 CGAATCAAGCTTATCGATACCGTCCAGCTCGAGGGGGGGCGCTACCCATTTT 1253
Db 1201 CGAATCAAGCTTATCGATACCGTCCAGCTCGAGGGGGGGCGCTACCCATTTT 1253

LOCUS AB094093 1202 bp mRNA linear PRI 02-MAY-2003
DEFINITION Homo sapiens DLNB14 mRNA, complete cds.
ACCESSION AB094093
VERSION AB094093.1 GI:30314480
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Kubo,T., Arai,Y., Ohira,M., Gamou,T., Maeno,G., Sakiyama,T., Toyoda,A., Hattori,M., Sakaki,Y., Nakagawara,A. and Ohki,M.
TITLE Identification of a 500-kb region of common allelic loss in chromosome 11q23 in non-MYC amplified type of neuroblastoma
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1202)
AUTHORS Arai,Y., Kubo,T., Ohira,M., Gamou,T., Maeno,G., Sakiyama,T., Toyoda,A., Hattori,M., Nakagawara,A. and Ohki,M.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2002) Yasuhiro Arai, National Cancer Center Research Institute, Molecular Oncology Division; 5-1-1 Teukiji, Chuo-ku, Tokyo 104-0045, Japan (E-mail: yaraig@n2.ncc.go.jp, Tel:81-3-3542-2511(ex.4752), Fax:81-3-3542-0688)
COMMENT This gene was identified within a LOH region at 11q23.3 of neuroblastoma. Reference sequence was predicted from genomic sequences of a BAC contig mapped at 11q23.3 using several gene prediction programs. Expression was analyzed by RT-PCR and Northern blot. Then, a part of this cDNA was obtained by cDNA library screening.
FEATURES
source location/Qualifiers
1..1202 /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
1..1202 /map="11q23.3"
1..1202 /gene="DLNB14"
57..1055 /gene="DLNB14"
/codon_start=1
/protein_id="BAC76047.1"
/db_xref="GI:30314481"
/translation="MAPAQCPCLRQTFPCGRHVYSRKHQRLKEALRLLPQVEAA RKFAAAQVRYVPEHEKCCWCLCGCEVRHLGNLTLYGLLEHLAGPEHKKAT KFWENKAEVQMEKEPLVTPQDYAPFKSMVKGLDSVEEKDKVTKMAAQIRVEQ SRQEVRSVLFPQAVPDEEGSSAPRSWKGNMSQVASSLOQPSNLDLPAPELDMET GPSLTFIIGHQDIPGVNIHSGATFPWMIQDEIYTAGNOEIGPSVEEFLEKEKQKLKK LPPDRVGANFHDHSRTSAGWLPSPFGRVWNNRRWQSRHOFKTEAAAMKKQSHTKS"
ORIGIN
Query Match 81.5%; Score 1021.8; DB 9; Length 1202;
Best Local Similarity 98.7%; Pred. No. 5.5e-260;
Matches 1093; Conservative 0; Mismatches 7; Indels 7; Gaps 6;
Qy 81 CTTCTTTCTGTGTCGCGC-GGCACGTTTACAGCCGCAAGCACCCAGCGGAGGAG 139
Db 94 CTTCTTTCTGTGTCGCGGACGTTTACAGCCGCAAGCA-CCAGCGGAGGAG 152
Qy 140 GCTTTTGTAGAGGCTCTGCCCCAGGTGAGGCGGCCGCAAGGCCATCGCGCCGCTCAG 199
Db 153 GCTTTGTAGAGGCTCTGCCCCAGGTGAGGCGGCCGCAAGGCCATCGCGCCGCTCAG 212
Qy 200 GTGGAGCGCTATGTGCCGGAACACAGCGATGCTCTGTCGCTCTGTCGGCTGTGAG 259
Db 213 GTGGAGCGCTATGTGCCGGAACACAGCGATGCTCTGTCGCTCTGTCGGCTGTGAG 272
Qy 260 GTGGCGGAACACCTGAGCCATGGAACCTGACGTCGTGTACGGGGGGCTGCTGGAGCAT 319
Db 273 GTGGCGGAACACCTGAGCCATGGAACCTGACGTCGTGTACGGGGGGCTGCTGGAGCAT 332
Qy 320 CTGGCCAGCCGAGCAGCAAGAAAGCAACAAATTTCTGTTGGGAGNACAAAGCTGAG 379

Db	333	CTGGCCAGCCAGACACAAAGAAACCAACCAAAATTTCTGGTGGGAAACAAGCTGAG	392
Qy	380	GTCCAGATGAAGAAGATTTCTGTCTACTCCCAAGGATTATGCGCGATTCAAGAAATCC	439
Db	393	GTCCAGATGAAGAAGATTTCTGTCTACTCCCAAGGATTATGCGCGATTCAAGAAATCC	452
Qy	440	ATGCTGAAGGTTTGATTTCTATGAAGAAGAGGATTAAGTATCAAGGAGATGGCA	499
Db	453	ATGCTGAAGGTTTGATTTCTATGAAGAAGAGGATTAAGTATCAAGGAGATGGCA	512
Qy	500	GCTCAGATCCGTGAGCTGAGCAGACGACGAGAGGTGGTTCGGTCTGTCTTAGAGCCT	559
Db	513	GCTCAGATCCGTGAGCTGAGCAGACGACGAGAGGTGGTTCGGTCTGTCTTAGAGCCT	572
Qy	560	CAGCAGTGCCAGACCCAGAGAGGGCTCTTCAGCACCTAGAACCTGGAAGGGATGAAC	619
Db	573	CAGCAGTGCCAGACCCAGAGAGGGCTCTTCAGCACCTAGAACCTGGAAGGGATGAAC	632
Qy	620	AGCCAAGTAGCTTCAGCTTACAGCAGCGCTCAAAATTTGGACCTGCCACAGCTCCAGAG	679
Db	633	AGCCAAGTAGCTTCAGCTTACAGCAGCGCTCAAAATTTGGACCTGCCACAGCTCCAGAG	692
Qy	680	CTTGACTCGATGGAGACAGGACCATCTCTGACATTTTCATTTGGCCATCAGGATATACCAGGA	739
Db	693	CTTGACTCGATGGAGACAGGACCATCTCTGACATTTTCATTTGGCCATCAGGATATACCAGGA	752
Qy	740	GTTGGTAATCACTCAGAGTGCCACACCTCCCTGGATGATCCAAAGATGAAGAATACATT	799
Db	753	GTTGGTAATCACTCAGAGTGCCACACCTCCCTGGATGATCCAAAGATGAAGAATACATT	812
Qy	800	GCTGGAAACCAAGAAATAGGACCATCTCTGACATTTTCATTTAAAGAAAGGAAAAACAG	859
Db	813	GCTGGAAACCAAGAAATAGGACCATCTCTGACATTTTCATTTAAAGAAAGGAAAAACAG	872
Qy	860	AAGTTGAAAAAATCCCTCCAGACGAGTGGGCGCACTTTGATCAGCTCCAGGACC	919
Db	873	AAGTTGAAAAAATCCCTCCAGACGAGTGGGCGCACTTTGATCAGCTCCAGGACC	932
Qy	920	AGTCAGGCTGGCTGCTCTTTTGGGCGCGCTCTGGAATAATGAGCCGCTGGCAGTC	979
Db	933	AGTCAGGCTGGCTGCTCTTTTGGGCGCGCTCTGGAATAATGAGCCGCTGGCAGTC	991
Qy	980	CAGACATCAATCCAAAATGAGCTGAGCAATGAAGAAGAGTTCATATACAGAAAAA	1039
Db	992	CAGACATCAA-TTCAAAAATGAGCTGAGCAATGAAGAAGAGTTCATATACAGAAAAA	1050
Qy	1040	GCTAATCATGCTCTACCAACTACCATGAGCTAAAGCCAAAGTCAACCAACCCCTA	1099
Db	1051	GCTAATCATGCTCTACCAACTACCATGAGCTAAAGCCAAAGTCAACCAACCCCTA	1108
Qy	1100	TTATACCTTCCACCCAAATTTCTTATCATTTCTTTCTTAGGAAACAGACATATCATTC	1159
Db	1109	TTATACCTTCCA-CCAAATTTCTTATCATTTCTTTCTTAGGAAACAGACATATCATTC	1167
Qy	1160	ATTGATTTAATAAGTTTATTTTC	1186
Db	1168	ATTGATTTAATAAGTTTATTTTC	1194
RESULT 3			
BC055929			
LOCUS			
DEFINITION	Mus musculus RIKEN cDNA D630044F24 gene, mRNA (cdna clone IMAGE:4205568), partial cds.	1175 bp	linear
ACCESSION	BC055929		
VERSION	BC055929.1	GI:33416549	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 1175) Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,		

TITLE	
JOURNAL	
PUBMED	
REFERENCE	
AUTHORS	
JOURNAL	
REMARK	
COMMENT	
FEATURES	
source	
gene	
CDS	
FEATURES	
Location/Qualifiers	
1..1175	
/organism="Mus musculus"	
/mol_type="mRNA"	
/strain="FVB/N"	
/db_xref="taxon:10090"	
/clone="IMAGE:4205568"	
/tissue_type="Colon, normal. 5 month old male mouse."	
/clone_lib="NCI CGAP_C024"	
/lab_host="DH10B"	
/note="vector: pCMV-SPORT6"	
<1..1175	
/gene="D630044F24R1k"	
/db_xref="LocusID:382073"	
/db_xref="MGI:2685960"	
<1..1018	
/gene="D630044F24R1k"	
/codon_start=2	
/product="D630044F24R1k protein"	
/protein_id="AAH55929.1"	
/db_xref="GI:33416550"	
/db_xref="LocusID:382073"	
/db_xref="MGI:2685960"	
/translation="FPGFSMAPQRCPLCRQTFCCRGHVYSHHQRLKGAERLLI PQVFAARRAQAQVRYVEPHDRCCWCPCCGCEVRKHLGNLTVLHGGLEHLASP EHKATNTKFWENKANAQMEKEFLISPDYARFKSMVKGLDSYEEKEDEVIKMAAQ	

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wyley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butcherfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1175)
Strausberg, R.
Direct Submission
Submitted (01-AUG-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 120 Row: j Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction.
Location/Qualifiers
1..1175
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4205568"
/tissue_type="Colon, normal. 5 month old male mouse."
/clone_lib="NCI CGAP_C024"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
<1..1175
/gene="D630044F24R1k"
/db_xref="LocusID:382073"
/db_xref="MGI:2685960"
<1..1018
/gene="D630044F24R1k"
/codon_start=2
/product="D630044F24R1k protein"
/protein_id="AAH55929.1"
/db_xref="GI:33416550"
/db_xref="LocusID:382073"
/db_xref="MGI:2685960"
/translation="FPGFSMAPQRCPLCRQTFCCRGHVYSHHQRLKGAERLLI
PQVFAARRAQAQVRYVEPHDRCCWCPCCGCEVRKHLGNLTVLHGGLEHLASP
EHKATNTKFWENKANAQMEKEFLISPDYARFKSMVKGLDSYEEKEDEVIKMAAQ


```
repeat_region 10. .329
repeat_region /note="MER87 repeat: matches 211. .540 of consensus"
repeat_region 544. .611
repeat_region /note="MERSA repeat: matches 121. .189 of consensus"
repeat_region 824. .852
repeat_region /note="MERSB repeat: matches 150. .178 of consensus"
repeat_region 865. .1291
repeat_region /note="L2 repeat: matches 307. .860 of consensus"
repeat_region 1565. .1604
repeat_region /note="10 copies 4 mer ttat 100% conserved"
repeat_region 1643. .2345
repeat_region /note="L1PB3 repeat: matches 5450. .6150 of consensus"
repeat_region 2534. .2687
repeat_region /note="L2 repeat: matches 833. .1002 of consensus"
repeat_region 2704. .3333
repeat_region /note="L2 repeat: matches 1774. .2440 of consensus"
repeat_region 3898. .4038
repeat_region /note="L1MB7 repeat: matches 5776. .5920 of consensus"
repeat_region 4108. .4463
repeat_region /note="L1MB7 repeat: matches 5303. .5661 of consensus"
repeat_region 4829. .5139
repeat_region /note="AlusX repeat: matches 4. .312 of consensus"
repeat_region 5208. .5388
repeat_region /note="AlusG/x repeat: matches 127. .309 of consensus"
repeat_region 5435. .5486
repeat_region /note="26 copies 2 mer tt 84% conserved"
repeat_region 5489. .5786
repeat_region /note="AlusX repeat: matches 1. .309 of consensus"
repeat_region 6844. .6959
repeat_region /note="58 copies 2 mer tt 62% conserved"
repeat_region 6978. .7364
repeat_region /note="L1MA9 repeat: matches 5894. .6275 of consensus"
repeat_region 7354. .7682
repeat_region /note="L1 repeat: matches 4721. .5050 of consensus"
repeat_region 7917. .8004
repeat_region /note="22 copies 4 mer ttcc 63% conserved"
repeat_region 8227. .8292
repeat_region /note="33 copies 2 mer ta 95% conserved"
repeat_region 8229. .8292
repeat_region /note="16 copies 4 mer tata 96% conserved"
repeat_region 10536. .10888
repeat_region /note="MER1B repeat: matches 1. .337 of consensus"
repeat_region 11927. .12045
repeat_region /note="L2 repeat: matches 2143. .2290 of consensus"
repeat_region 12303. .12811
repeat_region /note="MER75 repeat: matches 6. .514 of consensus"
gene 13443. .101318
CDS /gene="dJ718P11.1"
join(<13443. .13597,15536. .15684,20936. .>21054)
/note="match: ESTs: Em:AA192501"
/codon_start=1
/evidence=not experimental
/product="dJ718P11.1.2 (novel class II aminotransferase
similar to serine palmitoyltransferase (isoform 2))"
/protein_id="CAB65866.1"
/db_xref="GI:6691981"
/db_xref="GOA:Q9UGB5"
/db_xref="UniProt/TREMBL:Q9UGB5"
/translation="DFVPLYQDFNFYRNLYMRIRDNWRNRPICSAQPPLFLMERSV
DDTNWTFRTGRVKIDVINGSNFGLAAYKIDSMRTIKDVLVYGTGVASTRHEMG
EFTSSWQNALCVGVLQADSVQPPMSWQXGRHSSULE"
14113. .14582
/note="L1MA4 repeat: matches 5815. .6292 of consensus"
repeat_region 15902. .16080
repeat_region /note="L2 repeat: matches 2480. .2658 of consensus"
misc_feature complement(16079. .16250)
repeat_region /note="match: STS: Em:HSPF04H5"
18972. .19148
repeat_region /note="AluJo repeat: matches 121. .294 of consensus"
misc_feature 19894. .20269
repeat_region /gene="dJ718P11.1"
20181. .20417
repeat_region /note="match: GSS: Em:AQ044246"
20181. .20417
repeat_region /note="L1MA4 repeat: matches 4366. .4597 of consensus"
misc_feature complement(20538. .20941)
repeat_region /note="match: GSS: Em:AQ662868"
21219. .21378
repeat_region /note="L1MA4 repeat: matches 5160. .5326 of consensus"
repeat_region 21415. .21708
repeat_region /note="AlusC repeat: matches 1. .290 of consensus"
repeat_region 22230. .22469
repeat_region /note="L1R23 repeat: matches 206. .437 of consensus"
repeat_region 23617. .23863
repeat_region /note="MER58B repeat: matches 79. .337 of consensus"
repeat_region 24767. .24827
repeat_region /note="MIR repeat: matches 63. .123 of consensus"
repeat_region 25619. .25838
repeat_region /note="MIR repeat: matches 31. .238 of consensus"
repeat_region 25845. .26034
repeat_region /note="MIR repeat: matches 66. .258 of consensus"
misc_feature complement(25953. .26556)
repeat_region /note="match: GSS: Em:AQ482745"
27967. .28282
repeat_region /note="AluJo repeat: matches 1. .312 of consensus"
repeat_region 28751. .28859
repeat_region /note="L2 repeat: matches 2620. .2723 of consensus"
repeat_region 29512. .29809
repeat_region /note="AluY repeat: matches 1. .309 of consensus"
repeat_region 29810. .30116
repeat_region /note="AlusX repeat: matches 1. .296 of consensus"
repeat_region 30520. .30690
repeat_region /note="Charlie5 repeat: matches 2414. .2584 of consensus"
repeat_region 30691. .30996
repeat_region /note="AlusX repeat: matches 1. .306 of consensus"
repeat_region 30997. .31098
repeat_region /note="Charlie5 repeat: matches 2313. .2414 of consensus"
repeat_region 31097. .31278
repeat_region /note="Charlie5 repeat: matches -5. .115 of consensus"
repeat_region 31450. .31581
repeat_region /note="MIR repeat: matches 105. .249 of consensus"
misc_feature complement(31474. .32053)
repeat_region /note="match: GSS: Em:AQ083543"
31753. .32181
repeat_region /note="L1ME3 repeat: matches 5679. .6131 of consensus"
misc_feature 32059. .32525
repeat_region /gene="dJ718P11.1"
/note="match: GSS: Em:AQ774405"
Query Match 39.0%; Score 489; DB 9; Length.104913;
Best Local Similarity 82.3%; Pred. No. 3.7e-118;
```

Matches		672; Conservative	0; Mismatches	105; Indels	40; Gaps	8;
QY	406	CACTCCCCAGGATTATGCGGATTCAAGAAATCCATGGTCAAGGTTTGATTCCTATGA				
Db	15503					
QY	466	AGAAAAGGAGGATAAAGTGATCAAGGAGATGGCAGCTCAGATCCGTGAGGTGGAGCAGAG				
Db	15443					
QY	526	CCACAGGAGGTGGTTCGGTCTCTCTTAGAGCCT-----CAGGCAGTCC				
Db	15383	CTGGCAGGAGGTGGTTCAGTCTCTCTTAGAGGTGGTTCCTCGGAGGATCCAGACCCC				
QY	570	CAGACCCAGAGAGGCTCTTCAGCACCTAGAGCTGGAAAGGATGACAGCCAGTAG				
Db	15323	CTCAGCCAGTGC CGGCACTTTAGTCCCTCGGAAGCTGGAAAGGATGAACAGCCAGGTAG				
QY	630	CTTCCAGCTTACAGCAGCCCTCAAATTTGGACCTGCCACCAGCTCCAGAGCTTGACTGGA				
Db	15263	CTTCCAGCCACAGCAGGCGCTCCAACCTTGACCTGCCA-CAGCTCCAGAGCTTGACTGGA				
QY	690	TGGAGACAGACATCTCTGCACATTCATTGGCCATCAGATATACAGGAGTTGGTAACA				
Db	15204	TGGAGACAGGACTATCTCCGATATTCATTGGTCATCAGGATATACCAGGAGTTGGTAACA				
QY	750	TCCACTCAGGTGCCACACCTCCCTGGATGATCCAGATGAAGATACATTTGCTGGGAACC				
Db	15144	TCCACTCAGGTGCCACACCTCCCTGGATGATCCAGATGAAGATACATTTGCTGGGAACC				
QY	810	AAGAAATAGGACCATCTATGAAGATTTCTTAAAGAAAGGAAACACAGAGTTGMAAA				
Db	15084	AACAAACAGGACCGTCTCTATAAGATTTATTAAAGAAAGGAAAGAGAGTTGAAGG				
QY	870	AACTCCCCCAGACCGAGTTGGGCCAACTTTGATCACAAGCTCCAGGACATGCAGGCT				
Db	15024	AACTTTCCCAAACTAGTGTGGGCCAACTTTAATCACAGCTCCAGGACAGTGCAGGCT				
QY	930	GGCTGCCCTCTTTTGGGCGCGTCTGGAATAATGACGCGCTGGCAGTCCAGACATCAA				
Db	14964	GGCTGCCCTC-TTGGACCGTGTCTGGAATAATGATATCTGTGGCAGTTTCAGGCGTCAA				
QY	990	CTCCAAAATGAGCTGCAGCAATGAAGAAGCAGTCA-----CATACAGAAAAAAG				
Db	14905	CT-CAGAACTGAAGCTGCAGCAATGAACAAGCAGATAGCAGTCCATTTAAAAAAATG				
QY	1041	CTAATCATGC-----TCTCTACCACTACATGAGGCTAAAGC-----GAAATCAAC				
Db	14846	TTAATCATTTCCCTGATGCTCTACCAACTACCGTGAGGCTAAAGCAAAAGTCAACAACTC				
QY	1090	CAAAACCCCTATTATACCTTCCACCAGAAATCTTTATCATTTCTTTAGGAAACAGAC				
Db	14786	CTACCCACTATTATTCCTTCCA-CCAGTTCAATATCATTTGCTTTCTTAGGAAACAGAC				
QY	1150	ATACCTCATTTGATTATTAATAAGTTTATTTTTC				
Db	14727	ATGCCCATTCATTGATTATTAATAAGTTTATTTTCC				

RESULT 5
AC011089
LOCUS
DEFINITION
Homo sapiens chromosome 20 clone RP11-303K20 map 20, WORKING DRAFT
SEQUENCE, 13 unordered pieces.
AC011089
ACCESSION
AC011089.5
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
REFERENCE
1 (bases 1 to 192748)
Birken, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens chromosome 20, clone RP11-303K20

TITLE
JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 192748)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckerly, R., Boguslavskiy, L., Boukhgalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczký, J., Lieue, C., Locke, K., Macdonald, P., Marquis, N., Morrow, J., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (01-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 192748)

Birken, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F., Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Levine, R., Larocque, K., Lamazares, R., Landers, T., Lehoczký, J., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, D., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 3, 2000 this sequence version replaced gi:7717100.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1378

Center clone name: 303_K_20

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 182218 bases at least Q40

Consensus quality: 187129 bases at least Q30

Consensus quality: 189385 bases at least Q20

Insert size: 194000; agarose-fp

Insert size: 191548; sum-of-contigs

Quality coverage: 4.6 in Q20 bases; agarose-fp

Quality coverage: 4.7 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 13 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as


```
Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source           1..500
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
Query Match          35.5%; Score 444.4; DB 6; Length 500;
Best Local Similarity 97.8%; Pred. No. 1.1e-106;
Matches 493; Conservative 0; Mismatches 6; Indels 5; Gaps 4;
QY 683 GACTGATGAGACAGACACCATCTGACATTCATTGGCCATCAGGATATACAGAGTT 742
DB 1 GACTGATGAGACAGACACCATCTGACATTCATTGGCCATCAGGATATACAGAGTT 60
QY 743 GGTAAATCCTCAGTGGCCACACCTCCCTGGATGATCCAAAGATGAAGAATACATTGCT 802
DB 61 GGTAAATCCTCAGTGGCCACACCTCCCTGGATGATCCAAAGATGAAGAATACATTGCT 120
QY 803 GGGAAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAGAAAGGAAAAACAGAG 862
DB 121 GGGAAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAGAAAGGAAAAACAGAG 180
QY 863 TTGAAAAAATCTCCCCCAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCACT 922
DB 181 TTGAAAAAATCTCCCCCAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCACT 240
QY 923 GCAGGTGGTGGCCCTCTTTTGGCGCGCTCTGGAAATATGAACGCGCTGCGAGTCCAG 982
DB 241 GCAGGTGGTGGCCCTCTTTTGGCGCGCTCTGGAAATATGAACGCGCTGCGAGTCCAG 299
QY 983 ACATCAACTCCAAACCTGAAGCTGCAGCAATGAAGAGCAGTCACATACAGAAAAAGCT 1042
DB 300 ACATCAAA-TTCAAACTGAAGCTGCAGCAATGAAGAGCAGTCACATACAGAAAAAGCT 358
QY 1043 AATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCCAAAGTCAACAAACCCCTATTA 1102
DB 359 AATCATGCTCTCTACCAACTACCATGAGGCTAAAAGC--AAAGTCAACAAACCCCTATTA 416
QY 1103 TACCTCCACCAAAATCTTTTATCATGTCCTTTCTTAGGAAACAGACATACATTCATT 1162
DB 417 TACCTCCCA-CCAAATCTTTTATCATGTCCTTTCTTAGGAAACAGACATACATTCATT 475
QY 1163 TGATTTAATAAAGTTTATTTTTC 1186
DB 476 TGATTTAATAAAGTTTATTTTTC 499

RESULT 7
BV167722
LOCUS sqm6310 Human DNA (Sequenc) Homo sapiens STS 10-JUN-2004
DEFINITION tagged site.
ACCESSION BV167722
VERSION BV167722.1 GI:48001329
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 601)
Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
Cantor,C.R. and Braun,A.
Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
Regions
Genome Res. (2004) In press
Contact: Andreas Braun
Pharmaceuticals division
Sequenom, Inc.
3595 John Hopkins Court, San Diego, CA 92121, USA
Tel: 18582029018

BV167722
LOCUS sqm97502 Human DNA (Sequenc) Homo sapiens STS 10-JUN-2004
DEFINITION tagged site.
ACCESSION BV177918
VERSION BV177918.1 GI:48014136
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1191)
Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
Cantor,C.R. and Braun,A.
Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
Regions
Genome Res. (2004) In press
Contact: Andreas Braun
Pharmaceuticals division
Sequenom, Inc.
3595 John Hopkins Court, San Diego, CA 92121, USA
Tel: 18582029018
```

3595 John Hopkins Court, San Diego, CA 92121, USA

Tel: 18582029018

Fax: 18582029020

Email: abraun@sequenom.com

Primer A: No primer sequence submitted

Primer B: No primer sequence submitted

STS size: 1191.

Location/Qualifiers

1. .1191

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clone_lib="Human DNA (Sequenom)"

<1. .>1191

FEATURES

source

STS

ORIGIN

Query Match 31.9%; Score 400; DB 11; Length 1191;

Best Local Similarity 97.0%; Pred. No. 8.1e-95;

Matches 448; Conservative 1; Mismatches 8; Indels 5; Gaps 4;

Qy 725 CAGGATATACAGGAGTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAA 784

Db 270 CAGGATATACAGGAGTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAA 329

Qy 785 GATGAAGTAATACATTCCTGGGAACCAAGAAATAGGACCATCCTATGAAGAATTTCTTAAA 844

Db 330 GATGAAGTAATACATTCCTGGGAACCAAGAAATAGGACCATCCTATGAAGAATTTCTTAAA 389

Qy 845 GAAAGGAAACAGAGTTCGAAACCTCCACAGACCCAGTGGGGCCCACTTTGAT 904

Db 390 GAAAGGAAACAGAGTTCGAAACCTCCACAGACCCAGTGGGGCCCACTTTGAT 449

Qy 905 CACAGCTCCAGGACAGTGCAGGCTGGTGGCCCTCTTTTGGGCGCGTCTCGAATAATGG 964

Db 450 CACAGCTCCAGGACAGTGCAGGCTGGTGGCCCTCTTTTGGGCGCGTCTCGAATAATGG 508

Qy 965 ACGCGCTGGGAGTCCAGACATCAACTCCAAAACCTGAAGCTGCAGCAATGAAGACGAT 1024

Db 509 ACGCGCTGGGAGTCCAGACATCAAA-TTCAAAACCTGAAGCTGCAGCAATGAAGACGAT 567

Qy 1025 CACATACAGAAAAGCTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCCAAAG 1084

Db 568 CACATACAGAAAAGCTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCCAAAG 625

Qy 1085 TCAACCAACCCCTATTATACCTCCACCCAAATCTTTATCATTTCTTTCTTAGGAAA 1144

Db 626 GTCACAAACCCCTATTATACCTCCA-CCAAATCTTTATCATTTCTTTCTTAGGAAA 684

Qy 1145 CAGACATACCTATTCATTTGATTTAATAAGTTTATTTTTC 1186

Db 685 CAGACATACCTATTCATTTGATTTAATAAGTTTATTTTTC 726

RESULT 9

AX071560

LOCUS

DEFINITION

ACCESSION

AX071560

VERSION

AX071560.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

AX071560 401 bp DNA linear PAT 25-JAN-2001

Sequence 2032 from Patent WO0102568.

AX071560

AX071560.1 GI:12581911

Homo sapiens (human)

Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.,

Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,

Lamson, G., Drmanac, R., Crkencak, R., Drmanac, S., Dickson, M.,

Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and Strache-Crain, B.

Human genes and gene expression products

Patent: WO 0102568-A 2032 11-JAN-2001;

CHIRON CORPORATION (US); HYSEQ, INC. (US)

Location/Qualifiers

source

1. .401

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 31.3%; Score 392; DB 6; Length 401;

Best Local Similarity 98.8%; Pred. No. 9.1e-93;

Matches 395; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 475 GGATAAAGTATCAAGGAGATGGCAGCTCAGATCCGTGAGTGAGCAGAGCCGACAGGA 534

Db 2 GCACGAGTATCAAGGAGATGGCAGCTCATATCCGTGAGTGAGCAGAGCCGACAGGA 61

Qy 535 GGTGTTCCGTTCTCTTTAGAGCCTCAGGCGTCCAGACCCAGAGAAGGGCTCTTCAGC 594

Db 62 GGTGTTCCGTTCTCTTTAGAGCCTCAGGCGTCCAGACCCAGAGAAGGGCTCTTCAGC 121

Qy 595 ACCTAGAAGCTGGAAGGGATGAACAGCAAGTAGTTCAGCTTACAGAGCCCTCAAA 654

Db 122 ACCTAGAAGCTGGAAGGGATGAACAGCAAGTAGTTCAGCTTACAGAGCCCTCAAA 181

Qy 655 TTTGGACCTGCCACAGCTCCAGAGCTTGACTGGATGGAGACAGACCATTCTGACATT 714

Db 182 TTTGGACCTGCCACAGCTCCAGAGCTTGACTGGATGGAGACAGACCATTCTGACATT 241

Qy 715 CATTGGCCATCAGGATATACAGGAGTTGGTAAACATCCACTCAGGTGCCACACCTCCCTG 774

Db 242 CATTGGCCATCAGGATATACAGGAGTTGGTAAACATCCACTCAGGTGCCACACCTCCCTG 301

Qy 775 GATGATCAAGATGAAGAATACATTCGTGGGAACCAAGAAATAGGACCATCTATGAAGA 834

Db 302 GATGATCAAGATGAAGAATACATTCGTGGGAACCAAGAAATAGGACCATCTATGAAGA 361

Qy 835 ATTTCCTTAAAGAAAGGAAACAGAAAGTTGAAAAAATC 874

Db 362 ATTTCCTTAAAGAAAGGAAACAGAAAGTTGAAAAAATC 401

RESULT 10

CQ752826

LOCUS

DEFINITION

ACCESSION

CQ752826

VERSION

CQ752826.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CQ752826 486 bp DNA linear PAT 03-FEB-2004

Sequence 38760 from Patent WO02068579.

CQ752826

CQ752826.1 GI:42390079

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.

Kits, such as nucleic acid arrays, comprising a majority of

humanexons or transcripts, for detecting expression and other uses

thereof

Patent: WO 02068579-A 38760 06-SEP-2002;

PE Corporation (NY) (US)

Location/Qualifiers

1. .486

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

Query Match 26.1%; Score 327.4; DB 6; Length 486;

Best Local Similarity 83.0%; Pred. No. 1.3e-75;

Matches 404; Conservative 0; Mismatches 66; Indels 17; Gaps 2;

Qy 416 GATTATCGCGATTCAAGAAATCCATGCTGAAGGTTGGATTCTCTATGAAGAAAGGAG 475

Db 1 GATTATCGCGATTCAAGAAATCCATGCTGAAGGTTGGATTCTCTATGAAGAAAGGAG 60

Qy 476 GATAAAGTATCAAGGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGGAG 535

[illegible]

RESULT 12
LOCUS BD058555 421 bp DNA linear PAT 27-AUG-2002
DEFINITION Secreted expressed sequence tags (sESTs).
ACCESSION BD058555
VERSION BD058555.1 GI:22604161
KEYWORDS JP 2001519666-A/410.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 421)
AUTHORS Jacobs,K., McCoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D.,
Tracy,M., Spaulding,V. and Agostino,M.J.
TITLE Secreted expressed sequence tags (sESTs)
JOURNAL Patent: JP 2001519666-A 410 23-OCT-2001;
GENETICS INSTITUTE INC
COMMENT PN JP 2001519666-A/410
PD 23-OCT-2001
PF 10-APR-1998 JP 1998543068
PR 10-APR-1997 US 08/835913
PI KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE, PI
DAVID MERBERG,
PI MAURICE TRACY,VIKKI SPAULDING,MICHAEL J AGOSTINO PC
C12N15/12,C12N5/10,C07K14/47,C12Q1/68,A61K38/17 CC Strandedness:
Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source Location/Qualifiers
1..421
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 23.7%; Score 297; DB 6; Length 421;
Best Local Similarity 96.8%; Pred. No. 1.6e-67;
Matches 303; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 604 CTGGAAGGATGAACAGCCAGTAGCTTCAGCTTACAGAGCGCTCAAAATTTGGACCT 663
Db 100 CTGGCTCATGTTTTCAGCCAGTAGCTTCAGCTTACAGAGCGCTCAAAATTTGGACCT 159
Qy 664 GCCACAGCTCCAGAGCTTGACTGGATGGACAGACCATCTCTGCATTCATTGGCCA 723
Db 160 GCCACAGCTCCAGAGCTTGACTGGATGGACAGACCATCTCTGCATTCATTGGCCA 219
Qy 724 TCAGGATATACAGAGTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCA 783
Db 220 TCAGGATATACAGAGTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCA 279
Qy 784 AGATGAAGATACATTTGCTGGGAACCAAGAATAGGACCATCTATGAAGATTTCTTAA 843
Db 280 AGATGAAGATACATTTGCTGGGAACCAAGAATAGGACCATCTATGAAGATTTCTTAA 339
Qy 844 AGAAAGCAAAACACAGAGTTGAAATACTCCCCCAGACCGAGTTGGGGCCAACTTTGA 903
Db 340 AGAAAGCAAAACACAGAGTTGAAATACTCCCCCAGACCGAGTTGGGGCCAACTTTGA 399
Qy 904 TCACAGCTCCAGG 916
Db 400 TCACAGCTCCAG 412
RESULT 13
LOCUS CQ727443 433 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 13377 from Patent WO02068579.
ACCESSION CQ727443
VERSION CQ727443.1 GI:42293443
KEYWORDS Homo sapiens (human)
SOURCE

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 13377 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1..433
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 22.5%; Score 281.4; DB 6; Length 433;
Best Local Similarity 98.7%; Pred. No. 2.2e-63;
Matches 315; Conservative 0; Mismatches 1; Indels 3; Gaps 3;
Qy 727 GGATATACCGAGCTTGTAACTCCAC-TCAGGTGCCACACCTCCCTGGATGATCCAAG 785
Db 117 GGATATACCGAGCTTGTAACTCCAC-TCAGGTGCCACACCTCCCTGGATGATCCAAG 176
Qy 786 ATGAAGAATACTTGTCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAG 845
Db 177 ATGAAGAATACTTGTCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAG 236
Qy 846 AAAAGGAAAAACAGAAAGTTGAAAAAACTCCCCCAGACCGAGTTGGGGCCAACTTTGATC 905
Db 237 AAAAGGAAAAACAGAAAGTTGAAAAAACTCCCCCAGACCGAGTTGGGGCCAACTTTGATC 296
Qy 906 ACAGCTCCAGGACCGAGTGCAGGCTGGCTGCTCTTTGGGCGCGCTGGAATAATGGA 965
Db 297 ACAGCTCCAGGACCGAGTGCAGGCTGGCTGCTCTTTGGGCGCGCTGGAATAATGGA 355
Qy 966 CGCGCTGGAGTCCAGACATCAAACTCCAAACTGGAAGCTCAGCAATGAAGAAGCAGTC 1025
Db 356 CGCGCTGGAGTCCAGACATCAAACTCCAAACTGGAAGCTCAGCAATGAAGAAGCAGTC 414
Qy 1026 ACATACAGAAAAAGCTAA 1044
Db 415 ACATACAGAAAAAGCTAA 433
RESULT 14
LOCUS CQ728296 263 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 14230 from Patent WO02068579.
ACCESSION CQ728296
VERSION CQ728296.1 GI:42296746
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 14230 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1..263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 21.0%; Score 263; DB 6; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.6e-58;

Search completed: June 10, 2005, 14:22:55
Job time : 3569.59 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 22:58:14 ; Search time 425.87 Seconds
(without alignments)
17417.120 Million cell updates/sec

Title: US-09-155-676B-4
Perfect score: 1253
Sequence: 1 cattggagtcacgcggtggc.....ggggggccgtaccacctttt 1253

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002as: *
7: Geneseqn2002bs: *
8: Geneseqn2003as: *
9: Geneseqn2003bs: *
10: Geneseqn2003cs: *
11: Geneseqn2003ds: *
12: Geneseqn2004as: *
13: Geneseqn2004bs: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1252	99.9	1253	2	AAV03325 Sequence
2	514.6	41.1	681	5	AA83767 DNA encod
3	479.6	38.3	2557	5	AA83767 DNA encod
4	451.8	36.1	3948	6	AA83767 DNA encod
5	451.8	36.1	3948	10	AA83767 DNA encod
6	431.8	36.1	3948	12	AA83767 DNA encod
7	400.8	32.0	2556	2	AAV03325 Sequence
8	392	31.3	401	5	AA83767 DNA encod
9	370.4	29.6	742	5	AA83767 DNA encod
10	300.4	24.0	2489	9	AA83767 DNA encod
11	297	23.7	421	2	AAV03325 Sequence
12	263.4	21.0	3100	5	AA83767 DNA encod
13	244.8	19.5	374	3	AA83767 DNA encod
14	226	18.0	326	2	AAV03325 Sequence
15	222	17.7	584	10	AA83767 DNA encod
16	215.2	17.2	4017	5	AA83767 DNA encod
17	175.8	14.0	2803	4	AA83767 DNA encod
18	175.8	14.0	3731	4	AA83767 DNA encod
19	175.8	14.0	5269	4	AA83767 DNA encod
20	171	13.6	695	4	AA83767 DNA encod

21	140.6	11.2	214	4	AAK61910 Human imm
22	132.8	10.6	1600	5	AA83767 DNA encod
23	130.8	10.4	467	5	AA83767 DNA encod
24	122.8	9.8	328	3	AA83767 DNA encod
25	104.8	8.4	258	6	AA83767 DNA encod
26	83.4	6.7	97	3	AA83767 DNA encod
27	61.6	4.9	1906	2	AAV03323 5' end of
28	54.6	4.4	2000	8	AAV03323 5' end of
29	53.2	4.2	2631	2	AAV03324 Clone 10
30	46.8	3.7	2126	2	AAV03324 Clone 10
31	46.8	3.7	2126	2	AAV03324 Clone 10
32	46	3.7	10732	3	AAV03324 Clone 10
33	45.8	3.7	864	11	AAV03324 Clone 10
34	45	3.6	647	10	AAV03324 Clone 10
35	44.8	3.6	660	8	AAV03324 Clone 10
36	44.8	3.6	801	8	AAV03324 Clone 10
37	44.8	3.6	1236	2	AAV03324 Clone 10
38	44.8	3.6	1236	4	AAV03324 Clone 10
39	44.4	3.5	437	5	AAV03324 Clone 10
40	44.4	3.5	1889	4	AAV03324 Clone 10
41	43.8	3.5	1980	2	AAV03324 Clone 10
42	43.8	3.5	1980	2	AAV03324 Clone 10
43	43.2	3.4	468	5	AAV03324 Clone 10
44	43.2	3.4	468	5	AAV03324 Clone 10
45	43.2	3.4	478	5	AAV03324 Clone 10

ALIGNMENTS

RESULT 1
AAV03325
ID AAV03325 standard; cDNA; 1253 BP.
AC AAV03325;
DT 15-APR-1998 (first entry)
XX Sequence of clone 15, which encodes a TRAF2 binding protein.
DE Human tumour necrosis factor receptor-associated factor 2; TRAF2;
KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;
KW intracellular signalling activity; acute hepatitis;
KW autoimmune-induced cell death; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 2..1252
XX FT /*tag= a
XX FT /note= "no ATG start or STOP codon given"
XX
XX WO9737016-A1.
XX
XX 09-OCT-1997.
XX
XX 01-APR-1997; 97WO-IL000117.
XX
XX 02-APR-1996; 96IL-00117800.
XX 26-AUG-1996; 96IL-00119133.
XX (YEDA) YEDA RES & DEV CO LTD.
XX Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;
XX WPI; 1997-503101/46.
XX P-PSDB; AAW42401.
XX
XX DNA encoding tumour necrosis factor receptor-associated factor binding
XX molecule - used for modulation or mediation in cells of the activity of
XX NF-KB.
XX
XX Claim 4; Fig 5; 127pp; English.

XX The present sequence is that of clone 15, which encodes a TRAF2 binding
CC protein. Clone 15 is a partial clone, which lacks most of its 5' end of
CC the coding DNA sequence. A cDNA library prepared from B-cells was
CC screened for proteins that associate with TRAF2, and the present sequence
CC isolated. The clone 15 protein is capable of binding to at least amino
CC acids 222-501 of TRAF2. The TRAF-2 binding proteins can be used for
CC modulation or mediation in cells of the activity of NF-kappaB or any
CC other intracellular signalling activity modulated or mediated by TRAF2.
CC TRAF-binding proteins are especially used for prevention or treatment of
CC pathological conditions associated with NF-kB induction, e.g. acute
CC hepatitis, autoimmune-induced cell death, e.g. death of the beta
CC Langerhans cells or the pancreas that results in diabetes, the death of
CC cells in graft rejection, the death of oligodendrocytes in the brain in
CC multiple sclerosis, and AIDS-inhibited T cell suicide which causes
CC proliferation of the AIDS virus and hence the AIDS disease. The proteins
CC are also useful for screening of ligands capable of binding to a protein,
CC which are useful for modulating cellular activity modulated/mediated by
CC TRAF2
XX
SQ Sequence 1253 BP; 341 A; 321 C; 334 G; 256 T; 0 U; 1 Other;
Query Match 99.9%; Score 1252; DB 2; Length 1253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATTGGAGTCACGCGTGGCGGCTCTAGATAGTGGATCCCGGGCTGANGGAATTC 60
DB 1 CATTGGAGTCACGCGTGGCGGCTCTAGATAGTGGATCCCGGGCTGANGGAATTC 60
QY 61 GATTCAGCCAGCAAGGCCCTCTCTGTGGTGGCGCAGCTTTACAGCCGCAAGCAC 120
DB 61 GATTCAGCCAGCAAGGCCCTCTCTGTGGTGGCGCAGCTTTACAGCCGCAAGCAC 120
QY 121 CCAGCGCAGCTGAAGAGGCTTTTCAGAGGCTCTCTGCCCGCAGGTGGAGCGCGCCGCAA 180
DB 121 CCAGCGCAGCTGAAGAGGCTTTTCAGAGGCTCTCTGCCCGCAGGTGGAGCGCGCCGCAA 180
QY 181 GGCCATCCGCGCGCTCAGTGGAGCGCTATGTGCCCGCAACACAGCGATGCTGCTGGTG 240
DB 181 GGCCATCCGCGCGCTCAGTGGAGCGCTATGTGCCCGCAACACAGCGATGCTGCTGGTG 240
QY 241 CCTGTGCTGGCGCTGTGAGTGGCGGAAACACCTGAGCCATGGAAAACCTGACGGTGTGTA 300
DB 241 CCTGTGCTGGCGCTGTGAGTGGCGGAAACACCTGAGCCATGGAAAACCTGACGGTGTGTA 300
QY 301 CGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCAAGAAAGCAACAAATTCG 360
DB 301 CGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCAAGAAAGCAACAAATTCG 360
QY 361 GTGGGGAACAAAGCTGAGTCCAGATGAAGAGAGTTTCTGTCACCTCCCGAGGATTA 420
DB 361 GTGGGGAACAAAGCTGAGTCCAGATGAAGAGAGTTTCTGTCACCTCCCGAGGATTA 420
QY 421 TGCCCGATTCAAGAAATCCATGTTGAAGTTTGGATTCTCTATCAAGAAAGAGGATTA 480
DB 421 TGCCCGATTCAAGAAATCCATGTTGAAGTTTGGATTCTCTATCAAGAAAGAGGATTA 480
QY 481 AGTGATCAAGGAGATGGCAGCTCAGATCCGAGTGGAGAGCGCAGAGGAGTGGT 540
DB 481 AGTGATCAAGGAGATGGCAGCTCAGATCCGAGTGGAGAGCGCAGAGGAGTGGT 540
QY 541 TCGGTCTGTCTAGAGCTCAGGAGTGGCCAGAGCCAGAGAGGGCTCTTCAGCACCTAG 600
DB 541 TCGGTCTGTCTAGAGCTCAGGAGTGGCCAGAGCCAGAGAGGGCTCTTCAGCACCTAG 600
QY 601 AAGCTGGAAGAGGATGAACAGCAGTTCAGCTTCCAGCTTACAGAGCCCTCAAAATTTGGA 660
DB 601 AAGCTGGAAGAGGATGAACAGCAGTTCAGCTTCCAGCTTACAGAGCCCTCAAAATTTGGA 660
QY 661 CTTGCCACAGCTCCAGAGCTTGACTGTGATGGAGACAGGACCACTCTGACATTCATTGG 720
DB 661 CTTGCCACAGCTCCAGAGCTTGACTGTGATGGAGACAGGACCACTCTGACATTCATTGG 720

QY 721 CCATCAGGATATACAGGAGTTGGTAACATCCACTCAGTGGCCACACCTCCCTCGGATGAT 780
DB 721 CCATCAGGATATACAGGAGTTGGTAACATCCACTCAGTGGCCACACCTCCCTCGGATGAT 780
QY 781 CCAAGATGAAGAATACATTTCTCTGGGAACCAAGAAATAGGACCATCTATGAAGAATTTCT 840
DB 781 CCAAGATGAAGAATACATTTCTCTGGGAACCAAGAAATAGGACCATCTATGAAGAATTTCT 840
QY 841 TAAAGAAAAGGAAAAACAGAAAGTTGAAAAAATCTCCCCCAGACCGAGTTGGGGCCAACTT 900
DB 841 TAAAGAAAAGGAAAAACAGAAAGTTGAAAAAATCTCCCCCAGACCGAGTTGGGGCCAACTT 900
QY 901 TGATCAGAGTCCAGGACCAAGTGCAGGCTGGCTGGCTCTTTTGGGCGCGCTCTGGAATA 960
DB 901 TGATCAGAGTCCAGGACCAAGTGCAGGCTGGCTGGCTCTTTTGGGCGCGCTCTGGAATA 960
QY 961 ATGGACCCGCTGGCAGTCCAGACATCAACTCCAAAACTGAAAGCTGCAGCAATGAAGAAG 1020
DB 961 ATGGACCCGCTGGCAGTCCAGACATCAACTCCAAAACTGAAAGCTGCAGCAATGAAGAAG 1020
QY 1021 CAGTCACATACAGAAAAAGCTAATCATGCTCTACCAACTACCATGAGGCTAAAAAGCC 1080
DB 1021 CAGTCACATACAGAAAAAGCTAATCATGCTCTACCAACTACCATGAGGCTAAAAAGCC 1080
QY 1081 AAGTCAACCAAAACCCCTATTATACCTTCCACCAAAATCTTTATCATTTGTTCTTAG 1140
DB 1081 AAGTCAACCAAAACCCCTATTATACCTTCCACCAAAATCTTTATCATTTGTTCTTAG 1140
QY 1141 GAAACAGACATACATTCATTGATTAATAAGTTTATTTTCGGCCCTTCGTCGCT 1200
DB 1141 GAAACAGACATACATTCATTGATTAATAAGTTTATTTTCGGCCCTTCGTCGCT 1200
QY 1201 CGAATCAGCTTATCGATACCGTCAGCTCGAGGGGGGGCGGTACCCACATTTT 1253
DB 1201 CGAATCAGCTTATCGATACCGTCAGCTCGAGGGGGGGCGGTACCCACATTTT 1253
RESULT 2
AAS83767
ID AAS83767 standard; cDNA; 681 BP.
XX
AC AAS83767;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #19571.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSB-) HYSEQ INC.
XX
XX Dmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX P-PSDB; ABG19580.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX	PS	Claim 1; SEQ ID NO 19571; 103pp; English.	
XX	XX		
CC	CC	The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridization probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences	
XX	SQ	Sequence 681 BP; 154 A; 171 C; 231 G; 125 T; 0 U; 0 Other;	
		Query Match 41.1%; Score 514.6; DB 5; Length 681;	
		Best Local Similarity 97.4%; Pred. No. 2.6e-140;	
		Matches 568; Conservative 0; Mismatches 4; Indels 11; Gaps 4;	
Qy	81	CTTCTTTCTGTGTCGC-GGCACGCTTTACAGCCGCAAGCACCCAGCGAGCTGAAGGAG	139
Db	99	CTTCTTTCTGTGTCGCGGCGACGTTTACAGCGCGAAGCA-CCAGCGCGAGCTGAAGGAG	157
Qy	140	GCTTTTGAGAGCTTCCTGCCCC-----AGGTGGAGGCGCGCCCGAAGCCATCCGCG	191
Db	158	GCTTTTGAGAGGCTTCCTGCCCCAGGTGCGGAGGTGGAGCGCGCCCGAAGCCATCCGCG	217
Qy	192	CCGCTCAGGTGGAGCGCTATGTGCCGAACAGAGCGATGCTGCTGTGCTGTGCTGCG	251
Db	218	CCGCTCAGGTGGAGCGCTATGTGCCGAACAGAGCGATGCTGCTGTGCTGTGCTGCG	277
Qy	252	GCTGTAGGTGCGGGAACACCTGAGCCATGAAACCTGACGCTGCTGTACGGGGGCTGC	311
Db	278	GCTGTAGGTGCGGGAACACCTGAGCCATGAAACCTGACGCTGCTGTACGGGGGCTGC	337
Qy	312	TGGAGCATCTGGCCAGCCAGAGCACAAAGAACCAACAATTCGTGTGGAGAAC	371
Db	338	TGGAGCATCTGGCCAGCCAGAGCACAAAGAACCAACAATTCGTGTGGAGAAC	397
Qy	372	AAGCTGAGGTCCAGATGAAAGAGAGTTCTGTGTCACCTCCCGAGATTATGCGGATTCA	431
Db	398	AAGCTGAGGTCCAGATGAAAGAGAGTTCTGTGTCACCTCCCGAGATTATGCGGATTCA	457
Qy	432	AGAAATCCATGGTGAAGGTTTGGATTCCCTATGAAGAAAAGGAGGATTAAGTCAAGG	491
Db	458	AGAAATCCATGGTGAAGGTTTGGATTCCCTATGAAGAAAAGGAGGATTAAGTCAAGG	517
Qy	492	AGATGGAGCTCAGATCCGTGAGTGGAGCAGAGCCGACA-GGAGGTGGTTGGTCTGTCTC	550
Db	518	AGATGGAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAAGGAGGTGGTTGGTCTGTCTC	577
Qy	551	TTAGAGCTCAGGAGTCCAGAGCCAGAGAGGGCTTTCAGCACTTAGAGCTGGA	610
Db	578	TTAGAGCTCAGGAGTCCAGAGCCAGAGAGGGCTTTCAGCACTTAGAGCTGGA	637
Qy	611	GGGATGAACCCCAAGTAGCTTCCAGCTTACAGCAGCCCTCAA	653
Db	638	GGGATGAACCCCAAGTAGCTTCCAGCTTACAGCAGCCCTTAA	680
		RESULT 3	

AAS87242			
ID	AAS87242 standard; cDNA; 2557 BP.		
XX	AC	AAS87242;	
XX	DT	13-FEB-2002 (first entry)	
XX	DE	DNA encoding novel human diagnostic protein #23046.	
XX	KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;	
XX	KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.	
XX	OS	Homo sapiens.	
XX	PN	WO200175067-A2.	
XX	PD	11-OCT-2001.	
XX	PF	30-MAR-2001; 2001WO-US008631.	
XX	PR	31-MAR-2000; 2000US-00540217.	
XX	PR	23-AUG-2000; 2000US-00649167.	
XX	PA	(HYSE-) HYSEQ INC.	
XX	PI	Drmanac RT, Liu C, Tang YT;	
XX	DR	WPI; 2001-639362/73.	
XX	DR	P-PSDB; ABG23055.	
XX	PT	New isolated polynucleotide and encoded polypeptides, useful in	
XX	PT	diagnostics, forensics, gene mapping, identification of mutations	
XX	PT	responsible for genetic disorders or other traits and to assess	
XX	PT	biodiversity.	
XX	PS	Claim 1; SEQ ID NO 23046; 103pp; English.	
XX	CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)	
XX	CC	sequences. (I) is useful as hybridization probes, polymerase chain	
XX	CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,	
XX	CC	and in recombinant production of (II). The polynucleotides are also used	
XX	CC	in diagnostics as expressed sequence tags for identifying expressed	
XX	CC	genes. (I) is useful in gene therapy techniques to restore normal	
XX	CC	activity of (II) or to treat disease states involving (II). (II) is	
XX	CC	useful for generating antibodies against it, detecting or quantitating a	
XX	CC	polypeptide in tissue, as molecular weight markers and as a food	
XX	CC	supplement. (II) and its binding partners are useful in medical imaging	
XX	CC	of sites expressing (II). (I) and (II) are useful for treating disorders	
XX	CC	involving aberrant protein expression or biological activity. The	
XX	CC	polypeptide and polynucleotide sequences have applications in	
XX	CC	diagnostics, forensics, gene mapping, identification of mutations	
XX	CC	responsible for genetic disorders or other traits to assess biodiversity	
XX	CC	and to produce other types of data and products dependent on DNA and	
XX	CC	amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic	
XX	CC	coding sequences of the invention. Note: The sequence data for this	
XX	CC	patent did not appear in the printed specification, but was obtained in	
XX	CC	electronic format directly from WIPO at	
XX	CC	ftp.wipo.int/pub/published_pct_sequences	
XX	SQ	Sequence 2557 BP; 635 A; 624 C; 691 G; 606 T; 0 U; 1 Other;	
		Query Match 38.3%; Score 479.6; DB 5; Length 2557;	
		Best Local Similarity 84.9%; Pred. No. 1.1e-129;	
		Matches 702; Conservative 0; Mismatches 55; Indels 70; Gaps 12;	
Qy	426	GATTCAAGAAATCCATGGTGAAGGTTTGGATTCCTATGAAGAAAAGGAGGATTAAGTGA	485
Db	36	GATTCAAGAAATCCATGGTGAAGGTTTGGATTCCTATGAAGAAAAGGAGGATTAAGTGA	95
Qy	486	TCAAGGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGCCGACAGGAGGTGCTCGGT	545
Db	96	TCAAGGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGCCGACAGGAGGTGCTCGGT	155

```
QY 546 CTGTCTTAGAG-----CCTCAGCAGTGCACAGCCCA 577
Db |||||||
QY 156 CTGTCTTAGAGTTGGTTCCCTCGGAGATCCAGACCACTCAGGCGAGTGCCAGACCCA 215
Db |||||||
QY 578 GAAGAGGGCTCTTCAGCAGCTTA-GAAGCTGGAAAGGATGAACAGCCAAAGTAGCTT-CCA 635
Db |||||||
QY 216 GAAGAGGGCTCTTCAGCAGCTTAAGAGCTGGAAAGGATGAACAGCCAAAGTAGCTTCCA 275
Db |||||||
QY 636 GCTTACAGCAGCCCTCAAAATTTGGACCTGCCACAGCTCCAGAGCTTCACTGGATGGAGA 695
Db |||||||
QY 276 GCTTACAGCAGCCCTCAAAATTTGGACCTGCCACAGCTCCAGAGCTTCACTGGATGGAGA 335
Db |||||||
QY 696 CAGGACCACTCTGACATTCATTGGCCATCAGATATA-CCAGGAGTTGGTAACATCCAC 754
Db |||||||
QY 336 CAGGACCACTCTGACATTCATTGGCCATCAGATATAACCAGGAGTTGGTAACATCCAC 395
Db |||||||
QY 755 TCAGGTGCCACACCTCCCTGGATGATCCAGATGAAGAATACAT----TGCTGGGAACA 810
Db |||||||
QY 396 TCAGGTGCCACACCTCCCTGGATGATCCAGATGAAGAATACATTTGGCTGGGGGACCCCA 455
Db |||||||
QY 811 AGAATAGAGCC---ATCCTATGAAGATTTCTTAAGAAA----- 848
Db |||||||
QY 456 AGAATATNGGGCCCATCCCTATGAAGATTTCTTAAGAAAATTTCCCTATCATGTATC 515
Db |||||||
QY 849 --AGGAAAAACAGAGTTGAAAAAACTCCCCC---AGACCGAGTTGGGG--CCAACTT 900
Db |||||||
QY 516 AGAGGAAAAACAGAGTTGAAAAAACTCCCCC---AGACCGAGTTGGGG--CCAACTT 575
Db |||||||
QY 901 TGATCAGACCTCC-AGGACCAAGTGAGCTGGCTGCCCTCTTTTGGGGCGGCTCTGGAA 958
Db |||||||
QY 576 GGATCAGACCTCCAGGACCCAGTGCGAGCTGGCTGCCCTCTTTTGGGGCGGCTCTGGAA 635
Db |||||||
QY 959 TAATGACCGCGCTGGCAGTCCAGACATCAATCCAAACTGAAGTGCAGCAATGAAGA 1018
Db |||||||
QY 636 AATCGGACCGCGTGCGATCCAGACATCAATCCAAACTGAAGTGCAGCAATGAAGA 695
Db |||||||
QY 1019 AGCAGTCAATACAGAAAAAGCTAATCATGTCTCTACCAACTACCATGAGGCTAAAAG 1078
Db |||||||
QY 696 AGCAGTCAATACAGAAAAAGCTAATCATGTCTCTACCAACTACCATGAGGCTAAAAG 755
Db |||||||
QY 1079 CCAAAGTCAACCAACCCCTATATACCTTCCACCCCAAAATTTCTTATCATGTCTTTCT 1138
Db |||||||
QY 756 C--AAAGTCAACAAACCCCTATATACCTTCCA-CCAAATTTCTTATCATGTCTTTCTT 812
Db |||||||
QY 1139 AGGAAACAGACATCACTCATTTGATTTTGAATTAATAAGTTTATTTT 1185
Db |||||||
QY 813 AGGAAACAGACATCACTCATTTGATTTGAATTAATAAGTTTAAATTT 859
Db |||||||
```

```
RESULT 4
AAF88468
ID AAF88468 standard; cDNA; 3948 BP.
AC
XX
XX AAF88468;
DT 18-NOV-2002 (first entry)
XX
DE Human legless homologue hlgs-1 cDNA.
DE
XX
XX Legless; human; lgs; Wnt/Wingless signaling pathway; Wnt; Wg; gene;
KW tissue proliferation; tumour; cytostatic; cellular disorder; colon;
KW blood disorder; cancer; breast; head and neck cancer; brain; thyroid;
KW medulloblastoma; skin cancer; tissue regeneration; tissue repair; ss.
XX
OS Homo sapiens.
XX
XX Location/Qualifiers
FH Key 1..3948
FT CDS /*tag= a
FT /product= "hlgs-1"
FT misc_feature 601..3948
FT /*tag= b
FT /note= "region encoding protein represented in protein
```

```
FT AAB71230"
XX US2002086986-A1.
XX 04-JUL-2002.
XX 27-JUL-2001; 2001US-00915543.
XX 28-JUL-2000; 2000US-0221502P.
XX (BASL/) BASLER K.
XX (BRUN/) BRUNNER E.
XX (FROE/) FROESCH B.
XX (KRAM/) KRAMPS T.
XX (PETE/) PETER O.
XX Basler K, Brunner E, Froesch B, Kramps T, Peter O;
XX WPI; 2002-635689/68.
XX P-PSDB; ABB71230.
XX Novel polypeptide useful in therapeutic method for treating disorders of
XX cell fate such as cell differentiation or cell proliferation.
XX Claim 7; Fig 10A; 41pp; English.
XX This invention describes a novel polypeptide sharing one or more
XX homologous amino acid domains with the legless (lgs) protein, a
XX downstream component of the Wnt/Wingless (Wnt/Wg) signaling pathway
XX involved in the formation and maintenance of spatial arrangements and
XX proliferation of tissues during development, and in the formation and
XX growth of many human tumours. The products of the invention have
XX cytostatic activity and can be used to treat cellular disorders, blood
XX disorders and cancers caused by over-stimulation of the Wnt pathway,
XX where the cancerous condition is colon, breast, head and neck, brain,
XX thyroid, medulloblastoma or skin cancer. The product could also be used
XX to promote tissue regeneration and repair. This sequence encodes the
XX human legless (lgs) protein homologue hlgs-1 described in the disclosure
XX of the invention
XX Sequence 3948 BP; 829 A; 1272 C; 1220 G; 627 T; 0 U; 0 Other;
XX
XX Query Match 36.1%; Score 451.8; DB 6; Length 3948;
XX Best Local Similarity 99.2%; Pred. No. 2.1e-121;
XX Matches 475; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
QY 81 CTTCTCTTCTGTGTCGC-GGCACGTTTACAGCCGCAAGCACCCAGCGCAGCTGAAGGAG 139
Db |||||||
QY 3080 CTTCTCTTCTGTGTCGCAGCTTTACAGCCGCAAGCA-CCAGCGCAGCTGAAGGAG 3138
Db |||||||
QY 140 GCTTTTGAGAGGCTCTCTGCCCCCAGGTGGAGGCGGCCGCAAGGCATCCGCGCCCTCAG 199
Db |||||||
QY 3139 GCTTTTGAGAGGCTCTCTGCCCCCAGGTGGAGGCGGCCGCAAGGCATCCGCGCCCTCAG 3198
Db |||||||
QY 200 GTGGAGCGCTATGTGCCCCGAACACAGGCGATGCTGCTGTGCTGTGCTGCGGCTGTGAG 259
Db |||||||
QY 3199 GTGGAGCGCTATGTGCCCCGAACACAGGCGATGCTGCTGTGCTGTGCTGCGGCTGTGAG 3258
Db |||||||
QY 260 GTGCGGGAACACCTGAGCCATGGAACCTGACGGTGCTGTACGGGGGCTGCTGGAGCAT 319
Db |||||||
QY 3259 GTGCGGGAACACCTGAGCCATGGAACCTGACGGTGCTGTACGGGGGCTGCTGGAGCAT 3318
Db |||||||
QY 320 CTGGCCAGCCCCCAGAGACACAAAGAAAGCAACCAAAATTTCTGTTGGAGAACAAAGCTGAG 379
Db |||||||
QY 3319 CTGGCCAGCCCCCAGAGACACAAAGAAAGCAACCAAAATTTCTGTTGGAGAACAAAGCTGAG 3378
Db |||||||
QY 380 GTCCAGATGAAGAGAGAGTTTCTGTCTACTCCCGAGGATATGCCGGGCTTCAAGAAATCC 439
Db |||||||
QY 3379 GTCCAGATGAAGAGAGAGTTTCTGTGCTACTCCCGAGGATATGCCGGGCTTCAAGAAATCC 3438
Db |||||||
QY 440 ATGGGTGAAGAGTTTGTGATTTCTTATGAAGAAAGAGGAGTAAAGTGATCAAGGAGATGGCA 499
Db |||||||
QY 3439 ATGGGTGAAGAGTTTGTGATTTCTTATGAAGAAAGAGGAGTAAAGTGATCAAGGAGATGGCA 3498
Db |||||||
```


XX	Sequence	401 BP;	125 A;	94 C;	106 G;	76 T;	0 U;	0 Other;
Query Match	31.3%;	Score 392;	DB 5;	Length 401;				
Best Local Similarity	98.8%;	Pred. No. 2.2e-104;						
Matches 395;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;				
QY	475	GGATAAGTGATCAAGGAGATGGCAGCTCAGATCCGTGAGGTGGAGCGAGCGACGACGGA	534					
Db	2	GCACGAGGTGATCAAGGAGATGGCAGCTCATATCCGTGAGGTGGAGCGAGAGCCGACAGGA	61					
QY	535	GGTGGTTCGGTCTGTCTTAGAGGCTTCAGGCAGTGCAGACCCAGAGAGGGCTCTTCAGC	594					
Db	62	GGTGGTTCGGTCTGTCTTAGAGGCTTCAGGCAGTGCAGACCCAGAGAGGGCTCTTCAGC	121					
QY	595	ACCTAGAAGCTGGAAGGGATGAAACCCAAAGTAGCTTCCAGCTTACAGCAGGCCCTCAAA	654					
Db	122	ACCTAGAAGCTGGAAGGGATGAAACAGCCAAAGTAGCTTCCAGCTTACAGCAGGCCCTCAAA	181					
QY	655	TTTGGACCTGCCACGAGCTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATT	714					
Db	182	TTTGGACCTGCCACGAGCTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATT	241					
QY	715	CATTGGCCATCAGGATATACCCAGGAGTGTGTAACTCCACTCAGGTGCCACACCTCCCTG	774					
Db	242	CATTGGCCATCAGGATATACCCAGGAGTGTGTAACTCCACTCAGGTGCCACACCTCCCTG	301					
QY	775	GATGATCCAAGATGAAGAATACATCTCTGGGAACCAAGAAATAGGACCATCTCTATGAAGA	834					
Db	302	GATGATCCAAGATGAAGAATACATCTCTGGGAACCAAGAAATAGGACCATCTCTATGAAGA	361					
QY	835	ATTTCCTTAAAGAAAAGGAAAAACAGAGTTTGA AAAAATC	874					
Db	362	ATTTCCTTAAAGAAAAGGAAAAACAGAGTTTGA AAAAATC	401					

655	TTTGGACCTGCCACCAAGCTCCAGAGCTTGA	TGGAGACAGGACCAATCTCTGCACATT	714
182	TTTGGACCTGCCACCAAGCTCCAGAGCTTGA	TGGAGACAGGACCAATCTCTGCACATT	241
715	CATTGGCCATCAGGATATACCAAGGAGTTGG	TAACTCCACTCAAGTGCCACACTCCCTG	774
242	CATTGGCCATCAGGATATACCAAGGAGTTGG	TAACTCCACTCAAGTGCCACACTCCCTG	301
775	GATGATCCAAGATGAAGAATACATTCTCGGA	ACCAAGAAATAGGACCATCTTATGAAGA	834
302	GATGATCCAAGATGAAGAATACATTCTCGGA	ACCAAGAAATAGGACCATCTTATGAAGA	361
835	ATTTCCTTAAAGAAAGGAAAAACAGAAAGT	TGAAAAAATCTC	874
362	ATTTCCTTAAAGAAAGGAAAAACAGAAAGT	TGAAAAAATCTC	401

Qy	715	CATTGGCCCATCAGGATATACCAAGGAGTTGGTAACATCCACTCAGGTGCCACACCTCCCTG	774
Db	242	CATTGGCCCATCAGGATATACCAAGGAGTTGGTAACATCCACTCAGGTGCCACACCTCCCTG	301
Qy	775	GATGATCCAAGATGAAGAATACATCTCTGGGACCAAGAAATAGGACCATCTTATGAAGA	834
Db	302	GATGATCCAAGATGAAGAATACATCTCTGGGACCAAGAAATAGGACCATCTTATGAAGA	361
Qy	835	ATTCTTTAAAGAAAGGAAAAACAGAAAGTTGAAAAAATC	874
Db	362	ATTCTTTAAAGAAAGGAAAAACAGAAAGTTGAAAAAATC	401

242	CGTTGGCCATCAGGATATATACCCAGGAGTGGTGGATTCATCAGGATCCGACGATCCCTC	302
Qy		
775	GATGATCCAGATGAAGAATACATTCCTGGGACCAAGAAATAGGACCATCTCTATGAAGA	834
Qy		
302	GATGATCCAGATGAAGAATACATTCCTGGGACCAAGAAATAGGACCATCTCTATGAAGA	361
Db		
835	ATTCTTAAAGAAAGGAAAAACAGAGTTGAAAAAATCT	874
Qy		
362	ATTCTTAAAGAAAGGAAAAACAGAGTTGAAAAAATCT	401
Db		

QY	775	GATGATCCAGATGAAGAATACATCTCTGGGACCAAGAAATAGGACCATCTTATGAAGA	361
Db	302	GATGATCCAGATGAAGAATACATCTCTGGGACCAAGAAATAGGACCATCTTATGAAGA	361
QY	835	ATTTCCTTAAGAAAGGAAACACAGAGTTGAAAAAATC	874
Db	362	ATTTCCTTAAGAAAGGAAACACAGAGTTGAAAAAATC	401

[illegible]

Db
362 ATTCCTTAAGAAAAAGGAAAAACAGAAAGTTGAAAAAACTC 401

PS Claim 1; SEQ ID NO 19570; 103pp; English.


```
XX SQ Sequence 2489 BP; 699 A; 556 C; 506 G; 723 T; 0 U; 5 Other;
Query Match 24.0%; Score 300.4; DB 9; Length 2489;
Best Local Similarity 91.5%; Pred. No. 5.4e-77;
Matches 432; Conservative 0; Mismatches 21; Indels 19; Gaps 10;

Qy 715 CATTGGCCATCAGATATACAGAGATGGTGGTAAATCCATCCAGTGGCCACACCTCCCTG 774
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1969 CATTTCACCTACAGGATATAGGATGGTGGTAAATCCATCCAGTGGCCACACCTCCCTG 2028
Qy 775 GATGATCAAGATCAAGATATACATTCCTGGGAAACCAAGAAATAGGACCATCTATGAAGA 834
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2029 GATGATCAAGATGAAGA--ACATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGA 2086
Qy 835 ATTTCTTAAAGAAAGAAACAGAGATTTGAAATAAACTCCCCCCAGACCCAGTGGGGC 894
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2087 ATTTCTTAAAGAAAGG--AAACAGAGTTGAAATAAACTCCCCCCAGACCCAGTGGGGC 2144
Qy 895 CAACCTTGATCAGAGTCCAGGACGAGTGCGAGGCTGGCTGCTCTTTTGGGCGCGTCT 954
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2145 CAACCTTGATCAGAGTCC--GACCAGTGCGAGGCTGGCTGCTCTTTT-GGCGGTGCT 2201
Qy 955 GGAATAATGGACGCGCTGGCAGTCCAGACATCAACTCCAAACTGAAGCTGCAGCAATG 1014
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2202 GGAATAATGGACGCGCTGG--GTCCAGACATCAA-TTCAAACTGAAGCTGCAGCAATG 2258
Qy 1015 AAGAAGCAGTCAATACAGAAAGGCTTAATCATGCTCTCTPACCAACTACCATGAGGCTA 1074
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2259 AAGAAGCAGTCAATACAG--AAAGCTAATCATGCTCTCTACCACTACCATGAGGCTA 2316
Qy 1075 AAAGCCAAAGTCAACCAACCCCTATTATATACCTTCCACCCCAAAATCTTTATCATTTGCTT 1134
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2317 AAAG-CAAAGTCAACCAACCCCTTAT---ACCTTCCACCAAAATCTTTATCATTTGCTT 2371
Qy 1135 TCTTAGGAAACAGACATCACTCATTTGATTTTAAATAAGTTTATTTTC 1186
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2372 TCTTAGGAAACAGACATCACTCATTTGATTTTAAATAAGTTTATTTTC 2421

RESULT 11
AAV86432
ID AAV86432 standard; cDNA; 421 BP.
XX
AC AAV86432;
XX
DT 27-APR-1999 (first entry)
XX
DE EST clone AS20.
XX
KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW chemoraxis; chemokines; haemostasis; gene therapy; thrombolysis;
KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX
OS Homo sapiens.
XX
XX WO9845435-A2.
XX
XX 15-OCT-1998.
XX
XX 10-APR-1998; 98WO-US006954.
XX
XX 10-APR-1997; 97US-00835913.
XX
XX (GEMV ) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
XX Spaulding V, Agostino MJ;
XX
XX WPI; 1999-070076/06.
XX
XX New polynucleotides encoding human secreted proteins - derived from e.g.
```

```
PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,
XX pituitary, retina and colon cDNA libraries.
XX Claim 1; Page 232; 633pp; English.
XX
CC This sequence represents an expressed sequence tag (EST), and is a
CC polynucleotide of the invention. The polynucleotides of the invention are
CC all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity. The EST sequences are also stated to be useful for gene therapy
XX
SQ Sequence 421 BP; 128 A; 110 C; 88 G; 95 T; 0 U; 0 Other;
Query Match 23.7%; Score 297; DB 2; Length 421;
Best Local Similarity 96.8%; Pred. No. 2e-76;
Matches 303; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 604 CTGAAAGGATGAACAGCCAAAGTAGCTTCAGCTTACAGAGCCCTCAAAATTTGGACCT 663
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
100 CTGCTCATGTTTTCAGCCAAAGTAGCTTCAGCTTACAGAGCCCTCAAAATTTGGACCT 159
Qy 664 GCCACCAAGTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGGCCA 723
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
160 GCCACCAAGTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGGCCA 219
Qy 724 TCAGGATATACAGAGTGGTAACTCACTCAGGTGCCACACCTCCTCGATGATCCA 783
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
220 TCAGGATATACAGAGTGGTAACTCACTCAGGTGCCACACCTCCTCGATGATCCA 279
Qy 784 AGATGAAGAATACATTCCTGGGAAACCAAGAAATAGGACCATCTCTATGAAGAAATTTCTTAA 843
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
280 AGATGAAGAATACATTCCTGGGAAACCAAGAAATAGGACCATCTCTATGAAGAAATTTCTTAA 339
Qy 844 AGAAAGGAAACAGAGTTGAAATAAACTCCCCCAGACCGAGTTGGGGCCAACTTTGA 903
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
340 AGAAAGGAAACAGAGTTGAAATAAACTCCCCCAGACCGAGTTGGGGCCAACTTTGA 399
Qy 904 TCACAGCTCCAGG 916
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
400 TCACAGCTCCAG 412

RESULT 12
AAV77867
ID AAV77867 standard; cDNA; 3100 BP.
XX
AC AAV77867;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #13671.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
```


XX 22-SEP-1996 (first entry)
DE Human gene signature HUMGS06324.
XX
XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX
OS Homo sapiens.
XX
XX WO9514772-A1.
PN
XX
XX 01-JUN-1995.
PD
XX
XX 11-NOV-1994; 94WO-JP001916.
PF
XX
XX 12-NOV-1993; 93JP-00355504.
PR
XX
XX (MATS/) MATSUBARA K.
PA
XX (OKUB/) OKUBO K.
XX
XX Matsubara K, Okubo K; .
PI
XX WPI; 1995-206931/27.
DR
XX
XX Single-stranded DNA for identifying gene signatures - isolated from 3'-
PT directed human cDNA library that reflects relative abundance of corresp.
PT mRNA in specific human tissues.
XX
XX Claim 1; Page 1578; 2245pp; Japanese.
PS
XX
XX A single-stranded DNA (or its complementary strand or the corresp. double
CC -stranded DNA) which comprises one of the 7837 "GS" sequences given in
CC AAT19001-T26837 and which is able to hybridise to part of human genomic
CC DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences were
CC obtained from 3'-directed cDNA libraries prepared from various human
CC tissues; synthesis of cDNA was initiated from the 3'-end of mRNA by using
CC poly(T) as the sole primer. Since the 3'- untranslated sequence is unique
CC to a particular mRNA species, almost all the 3'-oriented cDNAs hybridise
CC with specific mRNAs. Each library is constructed so as to reflect
CC accurately the relative abundance of different mRNAs in the particular
CC tissue from which it was derived. The appearance frequency of a given GS
CC in a cDNA library can be determined (esp. using primers and probes
CC derived from the GS sequences) as a means of diagnosing abnormal cell
CC function or for recognising different cell types
XX
XX Sequence 326 BP; 87 A; 71 C; 64 G; 87 T; 0 U; 17 Other;

Query Match 18.0%; Score 226; DB 2; Length 326;
Best Local Similarity 95.8%; Pred. No. 1.3e-55;
Matches 273; Conservative 0; Mismatches 8; Indels 4; Gaps 4;
Qy 902 GATCACAGCTCCAGGACGAGCTGAGCTGGCTGCGCTCTTTTGGCGCGCTCGAATAA 961
Db 1 GATCACAGCTCCAGGACGAGCTGAGCTGGCTGCGCTCTTTTGGCGCGCTCGAATAA 59
Qy 962 TGGAGCGCGCTGGCAGTCCAGACATCACTCCAAAACCTGAAGCTGCAGCAATGAAGAC 1021
Db 60 TGGAGCGCGCTGGCAGTCCAGACATCACTCCAAAACCTGAAGCTGCAGCAATGAAGAC 118
Qy 1022 AGTCACATACAGAAAAAGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCCA 1081
Db 119 AGTCACATACAGAAAAAGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCCA 178
Qy 1082 AAGTCAACCAACCCCTATTATACCTTCCACCACCAATCTTTATCATGCTTCTTAGG 1141
Db 179 AAGTCAA-CAAAACCCCTATTATACCTTCCA-CCAAATCTTTATCATGCTTCTTAGG 236
Qy 1142 AAACAGACATCACTATTGATTTTAAATGAAGTTTATTTTC 1186
Db 237 AAACAGNCATCACTATTGATTTTAAATGAAGTTTATTTTC 281

RESULT 15
ADB49603/c
ID ADB49603 standard; DNA; 584 BP.
XX
XX ADB49603;
AC
XX
XX 04-DEC-2003 (first entry)
DT
XX
XX Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:145.
DE
XX
XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
XX Rattus norvegicus.
OS
XX
XX WO2003065993-A2.
PN
XX
XX 14-AUG-2003.
PD
XX
XX 04-FEB-2003; 2003WO-US003482.
PF
XX
XX 04-FEB-2002; 2002US-0353171P.
PR
XX 13-MAR-2002; 2002US-0363534P.
PR
XX 08-APR-2002; 2002US-0370248P.
PR
XX 10-APR-2002; 2002US-0371134P.
PR
XX 10-APR-2002; 2002US-0371135P.
PR
XX 10-APR-2002; 2002US-0371150P.
PR
XX 11-APR-2002; 2002US-0371413P.
PR
XX 19-APR-2002; 2002US-0373601P.
PR
XX 19-APR-2002; 2002US-0373602P.
PR
XX 22-APR-2002; 2002US-0374139P.
PR
XX 08-MAY-2002; 2002US-0378370P.
PR
XX 09-MAY-2002; 2002US-0378652P.
PR
XX 09-MAY-2002; 2002US-0378653P.
PR
XX 09-MAY-2002; 2002US-0378655P.
PR
XX 09-JUL-2002; 2002US-0394230P.
PR
XX 09-JUL-2002; 2002US-0394253P.
PR
XX 04-SEP-2002; 2002US-0407688P.
PR
XX 28-JAN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
FA
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
XX Elashoff M;
PI
XX
XX WPI; 2003-731472/69.
DR
XX
XX Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
XX Claim 44; SEQ ID NO 145; 874pp; English.
PS
XX
XX The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
XX Sequence 584 BP; 147 A; 121 C; 152 G; 164 T; 0 U; 0 Other;

Query Match		17.7%;	Score 222;	DB 10;	Length 584;
Best Local Similarity		76.0%;	Pred. No. 2.7e-54;		
Matches 327;		Conservative 0;	Mismatches 95;	Indels 8;	Gaps 4;
QY	757	AGGTGCGCACACCTCCCTGGATGATCCAAGATGAAGATACATTTGCTGGGAACCAAGAAAT	816		
Db	423	AGGTGCCACACCTCCATGGATGATCCAGAGGAGGAGCACAGCTCTGGAAGCCTCCCAAT	364		
QY	817	AGGACCATCCTATGAAGAAATTTCTTAAAGAAAAGGAAAAACAGAAAGTTGAAAAAACTCCC	876		
Db	363	AGGACCTCCTATGAAGAAATTTCTTAAAGAAAAGGAAAAACGAAAACTGAAGAAACTCCC	304		
QY	877	CCGAGACCAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCAGTCAGAGCTGGCTGCC	936		
Db	303	TCCAGATAGAGTTGGGGCCAACTTTGATCACAGTTCCAACACCACTGAGGCTGGCTGCC	244		
QY	937	CTCTTTTGGGCGCGTCTGGAATATGGACGCGCTGGCAGTCCAGACATCAACTCCAAA	996		
Db	243	CTCTTTT-GGCAGAGTCTGGAATATGGACGCGCTGGCAGTCCAGGCATCAA-TTCAAA	186		
QY	997	ACTGAAGCTGCAGCAATGAAGAAGCAGTCACATACAGAAAAAGCTAATCATGCTCTCTA	1056		
Db	185	ACTGAAGCTGCAGCAAGAGCAGGCAGCCACATAAAGGGAAAAAGCTGAAGTTCCCACTC	126		
QY	1057	CCAACTACCATGAGGCTAAAGCCAAAGTCAACCCAAACCCCTATTATACCTTCCACCCAA	1116		
Db	125	TGTTCCGATCTATGGGTAA-CAAAACCATTCAGCCAGCTATTATTCCTAAGACCAA	67		
QY	1117	ATTCTTTATCATGTCTTTCTTAGGAAACAGACATACCTCATTTGATTTTAATAAAGT	1176		
Db	66	-----TCCTCTCCAGCTGCCTGAGGAAACAGATGCACCCATTCATTTGATTCAATAAAGT	12		
QY	1177	TTTATTTTTC	1186		
Db	11	TTTATTTTCC	2		

Search completed: June 10, 2005, 06:10:33
Job time : 431.87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:20:34 ; Search time 2640.16 Seconds
(without alignments)
18065.048 Million cell updates/sec

Title: US-09-155-676B-4
Perfect score: 1253
Sequence: 1 cattggagtcacgcggtggc.....ggggggccgtaccacatttt 1253

Scoring table: IDENTITY NUC
Gapop 10*0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	755.6	60.3	926	2	BE796421 601592158
2	713.2	56.9	884	4	BI859740 603387142
3	694	55.4	1065	4	BQ403663 602419385
C 4	616.8	49.2	714	5	BQ008160 UI-H-E11-
5	587	46.8	816	4	BI769735 603055036
C 6	578.4	46.2	652	5	BQ07036 UI-H-E11-
C 7	576.4	46.0	758	5	BQ07036 UI-H-E11-
C 8	572	45.7	760	4	BM682005 UI-E-E01-
9	565	45.1	878	4	BI414452 602986067
10	563.8	45.0	652	7	CH291332 170006000
C 11	558.4	44.6	743	5	BQ608980 UI-CF-FN0
C 12	558.4	44.6	747	5	BQ608995 UI-CF-FN0
C 13	548.6	43.8	656	4	BM684305 UI-E-EJ1-
14	541.4	43.2	596	5	BM929870 UI-E-EJ1-
C 15	531.8	42.4	668	4	BI046266 MR3-FN020
C 16	530.8	42.4	870	3	AK052757 Mus muscu
C 17	519.4	41.5	599	6	CB321979 UI-CF-FN0
C 18	509	40.6	629	6	CA438500 UI-H-DT1-
C 19	507.4	40.5	706	5	BQ06283 UI-H-E11-
C 20	495.4	39.5	634	6	CA748930 UI-H-FE1-
C 21	486.4	38.8	680	5	BU733365 UI-E-CQ1-
C 22	486.4	38.8	759	5	BU733404 UI-E-CQ1-
C 23	486.2	38.8	790	6	CB989556 AGENCOURT
C 24	484.4	38.7	635	6	CA424880 UI-H-FE1-

C 25	456.8	36.5	520	2	BF446390
26	456	36.4	635	9	AY402903
27	448.4	35.8	548	2	BE669606
28	445.4	35.5	468	2	BE504243
29	442.4	35.3	635	9	AY402904
30	442.2	35.3	666	2	BE503633
31	437.4	34.9	704	7	CR454777
32	427	34.1	493	5	EX281470
33	426.2	34.0	571	2	AW957079
C 34	424.2	33.9	501	5	BM976229
C 35	420.6	33.6	534	2	AW243066
36	413.6	33.0	976	2	BF531630
37	403	32.2	735	7	CN792842
38	402.4	32.1	538	2	AW014772
39	400.8	32.0	502	2	BE222549
40	396	31.6	485	1	AV666223
41	394.2	31.5	748	4	BG701634
C 42	390.4	31.2	452	1	AI697110
C 43	382.8	30.6	444	4	BG231618
44	382.4	30.5	765	7	CO798321
45	380.2	30.3	497	2	BF077048

ALIGNMENTS

RESULT 1
BE796421
LOCUS BE796421 926 bp mRNA linear EST 20-SEP-2000
DEFINITION 601592158F1 NIH_MGC_7 Homo sapiens CDNA clone IMAGE:3945950 5',
mRNA sequence.
ACCESSION BE796421
VERSION BE796421.1 GI:10217619
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 926)
AUTHORS NTH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LNCM805 row: d column: 15
High quality sequence stop: 784.

FEATURES
source

Location/Qualifiers
1. 926
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3945950"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
ECORI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 60.3%; Score 755.6; DB 2; Length 926;

QY 859 GAAGTTGAAATACTCCCGCAGACCGAGTTGGGCGCAACTTTGATCAGCTCCAGAC 918
 DB |||||
 601 GAAGTTGAAATACTCCCGCAGACCGAGTTGGGCGCAACTTTGATCAGCTCCAGAC 660
 QY 919 CAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGCTCTGGAATAATGAGCG-CCGCTGGCAG 977
 DB |||||
 661 CAGTGCAGGCTGGCTGCCCTCTTTT-GGCGGCTCTGGAATAATGAGCGCGCTGGCAG 719
 QY 978 TCCAGACATCACTCAAACTGAAGCTGCAGCAATGAAGAGCAGTCAATACAGAAAA 1037
 DB |||||
 720 TCCAGACATCAA-TTCAAACTGAAGCTGCAGCAATGAAGAGCAGTCAATACAGAAAA 778
 QY 1038 AAGCTAATCATGCTCTTACCACTTACCATGAGGCTAAAGCCAAAGTCACCAACCCC 1097
 DB |||||
 779 AAG-TAATCATGCTCTTACCAATACCATGAGGCTAAAGCCAAAGTCACCAACCCC 837
 QY 1098 TATTATACCTTCCACCAAACTTTTATCATCTTCTTTTAGGAA 1143
 DB |||||
 838 TATTATACCTTCCACCAAACTTTTATCATCTTCTTTTAGGAA 883

RESULT 3
 BG403663
 LOCUS
 DEFINITION BG403663 1066 bp mRNA linear EST 12-MAR-2001
 mRNA sequence.

ACCESSION BG403663.1 GI:13297111

VERSION BG403663.

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE NIH-MGC <http://mgi.nci.nih.gov/>.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10433 row: j column: 21

High quality sequence stop: 735.

Location/Qualifiers

1..1066

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4526324"

/tissue type="transitional cell papilloma, cell line"

/lab host="DH10B (phage-resistant)"

/clone lib="NIH MGC 93"

/note="Organ: bladder; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dr primed.

Average insert size 1.7 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

FEATURES

source

RESULT 4

BQ006160/c

LOCUS

DEFINITION

UT-H-E11-ayy-j-22-0-UI.s1

IMAGE:5845365 3', mRNA sequence.

ACCESSION

BQ006160

VERSION

BQ006160.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 714)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Jose Mercuende

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be found

at:

<http://image.llnl.gov>

Plate: LLAM10433 row: j column: 21

High quality sequence stop: 735.

Location/Qualifiers

1..714

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5845365 3'

/tissue type="transitional cell papilloma, cell line"

/lab host="DH10B (phage-resistant)"

/clone lib="NIH MGC 93"

/note="Organ: bladder; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dr primed.

Average insert size 1.7 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

FEATURES

source

RESULT 4

BQ006160/c

LOCUS

DEFINITION

UT-H-E11-ayy-j-22-0-UI.s1

IMAGE:5845365 3', mRNA sequence.

ACCESSION

BQ006160

VERSION

BQ006160.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 714)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Jose Mercuende

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be found

at:

<http://image.llnl.gov>

Plate: LLAM10433 row: j column: 21

High quality sequence stop: 735.

Location/Qualifiers

1..714

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5845365 3'

/tissue type="transitional cell papilloma, cell line"

/lab host="DH10B (phage-resistant)"

/clone lib="NIH MGC 93"

/note="Organ: bladder; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dr primed.

Average insert size 1.7 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

FEATURES

source

RESULT 4

BQ006160/c

LOCUS

DEFINITION

UT-H-E11-ayy-j-22-0-UI.s1

IMAGE:5845365 3', mRNA sequence.

ACCESSION

BQ006160

VERSION

BQ006160.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 714)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Jose Mercuende

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be found

at:

<http://image.llnl.gov>

Plate: LLAM10433 row: j column: 21

High quality sequence stop: 735.

Location/Qualifiers

1..714

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5845365 3'

/tissue type="transitional cell papilloma, cell line"

/lab host="DH10B (phage-resistant)"

/clone lib="NIH MGC 93"

/note="Organ: bladder; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dr primed.

Average insert size 1.7 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

FEATURES

source

RESULT 4

BQ006160/c

LOCUS

DEFINITION

UT-H-E11-ayy-j-22-0-UI.s1

IMAGE:5845365 3', mRNA sequence.

ACCESSION

BQ006160

VERSION

BQ006160.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 714)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Jose Mercuende

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be found

at:

<http://image.llnl.gov>

Plate: LLAM10433 row: j column: 21

High quality sequence stop: 735.

Location/Qualifiers

1..714

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5845365 3'

/tissue type="transitional cell papilloma, cell line"

/lab host="DH10B (phage-resistant)"

/clone lib="NIH MGC 93"

/note="Organ: bladder; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dr primed.

Average insert size 1.7 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

FEATURES

source

RESULT 4

BQ006160/c

LOCUS

DEFINITION

UT-H-E11-ayy-j-22-0-UI.s1

IMAGE:5845365 3', mRNA sequence.

ACCESSION

BQ006160

VERSION

BQ006160.1

KEYWORDS

EST.

SOURCE

Homo sapiens

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
The following repetitive elements were found in this cDNA
sequence: 1-41, >AT rich/Low complexity
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

Location/Qualifiers
1..714
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5845365"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_E11"
/note="Organ: Left Pelvis; Vector: p7T73-Pac (Pharmacia)
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP E11 is a normalized cDNA library containing the
following tissue(s): Chondrosarcoma. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into p7T73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
ACACTTGCAC.
TAG TISSUE=chondrosarcoma
TAG LIB=UI-H-E11
TAG_SEQ=ACACTTGCAC"

ORIGIN

Query Match 49.2%; Score 616.8; DB 5; Length 714;
Best Local Similarity 98.1%; Pred. No. 1.2e-157;
Matches 677; Conservative 0; Mismatches 7; Indels 6; Gaps 5;
QY 497 GCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGGTTCGCTCTCTAGAG 556
DB 703 GCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGGTTCGCTCTCTAGAG 644
QY 557 CCTCAGCAGTGCAGACCCAGAGAGGGCTCTTCAGCACCTAGAGAGCTGGAAAGGGATG 616
DB 643 CCTCAGCAGTGCAGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 585
QY 617 AACAGCCAGTGTCTTCAGCTTACAGAGCCCTCAAAATTTGGACCTGCCACAGCTCCA 676
DB 584 AACAGCCAGTGTCTTCAGCTTACAGAGCCCTCAAAATTTGGACCTGCCACAGCTCCA 525
QY 677 GAGCTTGAATGGAG 736
DB 524 GAGCTTGAATGGAG 465
QY 737 GAGTTCGTGAATCCACTCAGTGCACACCTCCCTGGATGATCCTCAAGATGAAGATAC 796
DB 464 GGAGTTGGTGAATCCACTCAGTGCACACCTCCCTGGATGATCCTCAAGATGAAGATAC 405
QY 797 ATTGCTGGGAAACCAAGAAATAGGACCAATCTATGAAGAATTTCTTAAAGAAAGGAAAAA 856
DB 404 ATTGCTGGGAAACCAAGAAATAGGACCAATCTATGAAGAATTTCTTAAAGAAAGGAAAAA 345
QY 857 CAGAGCTTGAAGAAACCTCCCGCAGAGCCGAGTGGGGCCAACTTGTATACACCTCCAGG 916
DB 344 CAGAGCTTGAAGAAACCTCCCGCAGAGCCGAGTGGGGCCAACTTGTATACACCTCCAGG 285
QY 917 ACCAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGTCTTGAATTAATGGACGCCCTGGCA 976
DB 284 ACCAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGTCTTGAATTAATGGACGCCCTGGCA 226
QY 977 GTCCAGACATCAATCCAAAACCTGAAGCTGAGCAATGAAGAGAGAGTGCATACATACAGAAA 1036
DB 225 GTCCAGACATCAATCCAAAACCTGAAGCTGAGCAATGAAGAGAGAGTGCATACATACAGAAA 167

QY 1037 AAAGCTAATCATGCTCTCTCTACCACTACCATGAGGCTAAAGCCAAAGTCAACCAACCC 1096
DB 166 AAAGCTAATCATGCTCTCTCTACCACTACCATGAGGCTAAAGC-AAAGTCAACCAACCC 109
QY 1097 CTATTATACCTTCACCCAAATTTCTTTATCATTTGCTTTCTTAGGAAACAGACATATCA 1156
DB 108 CTATTATACCTTCACCCAAATTTCTTTATCATTTGCTTTCTTAGGAAACAGACATATCA 50
QY 1157 TTCAATTTGATTTAATAAAGTTTATTTTC 1186
DB 49 TTCAATTTGATTTAATAAAGTTTATTTTC 20
RESULT 5
BI769735 816 bp mRNA linear EST 25-SEP-2001
LOCUS 601055036F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5204408 5',
DEFINITION mRNA sequence.
ACCESSION BI769735
VERSION BI769735.1 GI:15761300
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 816)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11512 row: h column: 09
High quality sequence stop: 790.
Location/Qualifiers
1..816
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5204408"
/lab_host="DH10B"
/clone_lib="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb. Insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Query Match 46.8%; Score 587; DB 4; Length 816;
Best Local Similarity 91.6%; Pred. No. 1.8e-149;
Matches 721; Conservative 0; Mismatches 5; Indels 61; Gaps 7;
QY 81 CCTTCTTCTGTCGTCGC-GGCACAGTTTACAGCCGCAAGCACCAGCGGAGCTGAAGGAG 139
DB 58 CCTTCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 116
QY 140 GCTTTTGAAGAGCTCTCTGCCCCCAGGTGGAGCGCGCCGCAAGCCATCGCGCCGCTCAG 199
DB 117 GCTTTTGAAGAGCTCTCTGCCCCCAGGTGGAGCGCGCCGCAAGCCATCGCGCCGCTCAG 176


```

Qy 200 GTGAGCGCTATGTCCCGAACACAGCGATGCTCTGCTGCTGTGTGCTGGCGTGTGAG 259
Db 177 GTGGAGCGCTATGTGCGCGAACACAGCGATGCTCTGCTGCTGTGTGCTGGCGTGTGAG 236
Qy 260 GTGCGGGAACACCTGAGCCATGGAACCTGACGGTGTGACGGGGGCTGCTGGAGCAT 319
Db 237 GTGCGGGAACACCTGAGCCATGGAACCTGACGGTGTGACGGGGGCTGCTGGAGCAT 296
Qy 320 GTGCGGAGCCAGAGCACAGAAACCAACCAAAATTTCTGTTGGGGAACAAAGCTGAG 379
Db 297 GTGCGGAGCCAGAGCACAGAAACCAACCAAAATTTCTGTTGGGGAACCAAGCTGAG 356
Qy 380 GTCCAGATGAAGAGAGATTTCTGTCTACTCCAGGATTTATGCGCGATTCAAGAAATCC 439
Db 357 GTCCAGATGAAGAGAGATTTCTGTCTACTCCAGGATTTATGCGCGATTCAAGAAATCC 416
Qy 440 ATGGTGAA-AGCTTTGGATTCTATGAAGA--AAAGGAGGATAAGTATCATCAAGGAGATG 496
Db 417 ATGGTGAACAGGTTTGGATTCTATGAAGAACACAGGAGGATAAGTATCATCAAGGAGATG 476
Qy 497 GCAGTCTAGATCCGTGAGGTGGAGCAGAGCCGACAGAGGAGTGGTTCGGTCTGTCTTAGAG 556
Db 477 GCAGTCTAGATCCGTGAGGTGGAGCAGAGCCGACAGAGGAGTGGTTCGGTCTGTCTTAGAG 536
Qy 557 -----CCTCAGCAGTGCACAGACCAGAGAGGGGCTC 588
Db 537 GTTGGTTTCCCTCGGAGGATCCAGACCACTCAGGCAGTGCAGACCAGAGAGGGGCTC 596
Qy 589 TTCAGCACTAGAGCTGGAAGGATGAACAGACCCAGTACTTCCAGCTTACAGCAGCC 648
Db 597 TTCAGCACTAGAGCTGGAAGGATGAACAGACCCAGTACTTCCAGCTTACAGCAGCC 656
Qy 649 CTCAAAATTGGACCTGCCACAGCTTCCAGAGCTTGACTTGGGA-TGGAGACAGGACCATCTC 707
Db 657 CTCAAAATTGGACCTGCCACAGCTTCCAGAGCTTGACTTGGGA-TGGAGACAGGACCATCTC 716
Qy 708 TGACATTCATTGGCCATCAGGATATACAGAGTGGTAAATCATCACTCAGGTGCCACAC 767
Db 717 TGACATTCATTGGCCATCAGGATATACAGAGTGGTAAATCATCACTC----- 765
Qy 768 CTCCTCGATCCAGATCAGATATACATGCTGGGAACCAAGTAAGACATCCT 827
Db 766 -----AGATGAAGATATACATGCTGGGAACCAAGATAGGACCATCCT 809

Qy 828 ATGAAGA 834
Db 810 ATGAAGA 816

RESULT 6
BM972124/c
LOCUS
DEFINITION UI-CF-EC1-abp-m-17-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone
UI-CF-EC1-abp-m-17-0-UI 3', mRNA sequence.
ACCESSION BM972124
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 652)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866

```

Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com) or from Open Biosystems
 (www.openbiosystems.com).
 The following repetitive elements were found in this cDNA
 sequence: 1-41, >AT rich#Low_complexity
 Seq primer: M13 FORWARD
 POLYA=yes.

FEATURES

Location/Qualifiers
 1..652
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="UI-CF-EC1-abp-m-17-0-UI"
 /tissue_type="Lung"
 /dev_stage="Adult and Fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-CF-EC1"
 /notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-CF-EC1 is a normalized cDNA library containing the
 following tissue(s): Normal lung from adult and from fetal
 day 64, day 87, week 19 and week 42. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT7T3-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 AAGTGGCTTAC.
 TAG TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
 and 380-383
 TAG_LTB=UI-CF-EC1
 TAG_LBO=AAGTGGCTTAC"

ORIGIN

Query Match 46.2%; Score 578.4; DB 5; Length 652;
 Best Local Similarity 98.3%; Pred. No. 4e-147;
 Matches 627; Conservative 0; Mismatches 6; Indels 5; Gaps 4;

```

Qy 549 TCTTAGAGCCTCAGGCGAGTGCACAGACCAGAGAGGGCTCTTCAGCACCTAGAGCTGGA 608
Db 652 TCTTAGAGCCTCAGGCGAGTGCACAGACCAGAGAGGGCTCTTCAGCACCTAGAGCTGGA 593
Qy 609 AAGGATGAACAGCAAGTAGCTTCAGCTTACAGAGCCCTCAAAATTGGACCTGCCAC 668
Db 592 AAGGATGAACAGCAAGTAGCTTCAGCTTACAGAGCCCTCAAAATTGGACCTGCCAC 533
Qy 669 CAGCTCCAGAGCTTGACTGGATGGAGACAGACCATCTCTGACATTCATTGGCCATCAGG 728
Db 532 CAGCTCCAGAGCTTGACTGGATGGAGACAGACCATCTCTGACATTCATTGGCCATCAGG 473
Qy 729 ATATACCAGGAGTTGGTAAACATCCACTCAGGTGCCACACCTCCCTGGATGATCAAGATG 788
Db 472 ATATACCAGGAGTTGGTAAACATCCACTCAGGTGCCACACCTCCCTGGATGATCAAGATG 413
Qy 789 AAGAATACATGTGGGAACCAAGAAATAGGACATCCTATGAAGAAATTTCTTAAAGAAA 848
Db 412 AAGAATACATGTGGGAACCAAGAAATAGGACATCCTATGAAGAAATTTCTTAAAGAAA 353
Qy 849 AGGAAAACAGAGTTGAAAATACTCCCCAGACAGAGTTGGGGCCAACTTTGATCACA 908
Db 352 AGGAAAACAGAGTTGAAAATACTCCCCAGACAGAGTTGGGGCCAACTTTGATCACA 293
Qy 909 GCTCCAGGACAGTGCAGGCTGGCTCTTTTGGGCGCGTCTGGAATAATGAGCGC 968

```

|||||
292 GCTCCAGGACAGTGGCTGCTCCCTCTTTT-GGCGCGCTCTGGAATATGAGCGC 234
QY 969 CGCTGGCAGTCCAGACATCAACTCCAAAAGTGAAGTGGAGGATGAAGAGCAGTCA 1028
Db 233 CGCTGGCAGTCCAGACATCAA-ATCAAAAGTGAAGTGGAGGATGAAGAGCAGTCA 175
QY 1029 TACAGAAAAAGTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCCAAAGTCAA 1088
Db 174 TACAGAAAAAGTATCATGCTCTCTACCAACTACCATGAGGCTAAAAGC--AAAGTCA 117
QY 1089 CMAACCCCTATATATACCTTCCACCACCAATCTTTATCATGCTTTCTTAGGAAACAGA 1148
Db 116 ACAAAACCCCTATATATACCTTCCA-CCAAATCTTTATCATGCTTTCTTAGGAAACAGA 58
QY 1149 CATACTCATTCATTTGATTAATAAGTTTATTTTTC 1186
Db 57 CATACTCATTCATTTGATTAATAAGTTTATTTTTC 20

RESULT 7
BQ007036/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
BQ007036 758 bp mRNA linear EST 26-MAR-2002
UI-H-EII-azb-p-24-0-UI.s1 NCI_CGAP_EII Homo sapiens cDNA clone
IMAGE:5846663 3', mRNA sequence.
BQ007036
BQ007036.1 GI:19731936
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 758)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
CDNA library prepared by: Dr. M. Bento Soares, University of Iowa
CDNA library arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
The following repetitive elements were found in this cDNA
sequence: 1-41, sAT-rich#Low_complexity
Seq primer: 1-41, sAT-rich#Low_complexity
POLYA=Yes.

FEATURES
source
1. .758
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5846663"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_EII"
/note="Organ: Left Pelvis; Vector: pTT3-Pac (Pharmacia)
with a modified polylinker; Site 1: Ecor I; Site 2: Not I;
NCI_CGAP_EII is a normalized cDNA library containing the
following tissue(s): Chondrosarcoma. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an Ecor I
adaptor, digested with Not I, and cloned directionally
into pTT3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
ACACTTGCAC.

TAG TISSUE=chondrosarcoma
TAG_LIB=UI-H-EII
TAG_SEQ=ACACTTGCAC"
ORIGIN
Query Match 46.0%; Score 576.4; DB 5; Length 758;
Best Local Similarity 94.4%; Pred. No. 1.5e-146;
Matches 706; Conservative 0; Mismatches 6; Indels 36; Gaps 9;
QY 466 AGAAAAGAGGATAAAGTATCAAGAGATGGCAGCTCAGATCCGTGAGGTGGAGCAGAG 525
Db 758 AGAAAAGAGGATAAAGTATCAAGAGAT--GCAGCTCAGATCCGTGA--GTGAGCAGAG 701
QY 526 CCGACAGAGGTGGTTCGCTCTCTTAGAGGTTTCCCTCGGAGGATCCAGACCACC 558
Db 700 CCGACAGAGGT-GTTCGCTCTCTTAGAGGTTTCCCTCGGAGGATCCAGACCACC 642
QY 559 TCAGGCAGTCCAGACCCAGAGAGGGCTCTTTCAGCAGCTAGAAAGCTGGAAGGGATGAA 618
Db 641 TCAGGCAGTCCAGACCCAGAGAG--GGCTCTTCAGCAGCTAGAAAGCTGGAAGGGATGNA 583
QY 619 CAGCCAAAGTAGTCTTCAGCTTACAGCAGCCCTCAAATTTGGACCTGGCCACCGACTCCAGA 678
Db 582 CAGCCAAAGTAGTCTTCAGCTTACAGCAGCCCTCAAATTTGGACCTGGCCACCGACTCCAGA 523
QY 679 GCTTGACTGGATGGAGAGCAGGACCATCTCTGACATTCATTTGGCCATCAGGATATACCAGG 738
Db 522 GCTTGACTGGATGGAGAGCAGGACCATCTCTGACATTCATTTGGCCATCAGGATATACCAGG 463
QY 739 AGTTGGTAAACATCCACTCAGGTGGCCACACCTCCCTGGATGATCCAAAGATGAAGATACAT 798
Db 462 AGTTGGTAAACATCCACTCAGGTGGCCACACCTCCCTGGATGATCCAAAGATGAAGATACAT 403
QY 799 TGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAGAAAAAGGAAAAACA 858
Db 402 TGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAGAAAAAGGAAAAACA 343
QY 859 GAAGTTGAAAAAAGCTCCGCCAGACCGAGTTGGGCCCAACTTTGATACACAGTCCAGGAC 918
Db 342 GAAGTTGAAAAAAGCTCCGCCAGACCGAGTTGGGCCCAACTTTGATACACAGTCCAGGAC 283
QY 919 CAGTGCAGGCTGGCTGGCTCTTTTGGGCCGCGCTTGGAAATATGAGACCGCTGGCAGT 978
Db 282 CAGTGCAGGCTGGCTGGCTCTTTTGGGCCGCGCTCTTTT-GGCCGCGCTCTGGAATATGAGAC 224
QY 979 CCAGACATCAACTCCAAAAGTGAAGCTGCAGCAATGAAGAGCAGTCAATACATCAGAAAAA 1038
Db 223 CCAGACATCAA-TTCAAAAGTGAAGCTGCAGCAATGAAGAGCAGTCAATACATCAGAAAAA 165
QY 1039 AGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCCAAAGTCAACCAACCCCT 1098
Db 164 AGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGC--AAAGTCAACCAACCCCT 107
QY 1099 ATTATACCTTCCACCCCAAAATTTTATCATTTCTTTCTTAGGAAACAGACATCTCATT 1158
Db 106 ATTATACCTTCCA-CCAAATTTCTTTATCATTTCTTTCTTAGGAAACAGACATCTCATT 48
QY 1159 CATTTGATTTTAAAGTTTATTTTTC 1186
Db 47 CATTTGATTTTAAAGTTTATTTTTC 20
RESULT 8
BQ007036/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BQ007036 760 bp mRNA linear EST 27-FEB-2002
UI-E-E01-aiw-c-24-0-UI.s1 UI-E-E01 Homo sapiens cDNA clone
UI-E-E01-aiw-c-24-0-UI 3', mRNA sequence.
BQ007036
BQ007036.1 GI:18991901
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 760)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 1-41, >AT rich#Low_complexity
Seq primer: M13 Forward
POLYA=yes.

FEATURES
source

Location/Qualifiers
1..760
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EOI-a1w-c-24-0-UI"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EOI"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-EOI is a normalized cDNA library containing the
following tissue(s): fetal eye. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
CGCTATACC. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye
Institute (NEI).
TAG_TISSUE=human fetal eye
TAG_LIB=UI-E-EOI
TAG_SEQ=CGCTATACC"

ORIGIN

Query Match 45.7%; Score 572; DB 4; Length 760;
Best Local Similarity 93.9%; Pred. No. 2.4e-145;
Matches 704; Conservative 0; Mismatches 10; Indels 36; Gaps 9;

Qy 464 GAAGAAAGGAGGATAAAGTGATCAAGAGATGGCAGCTCAGATCCGTGGAGTGGAGCAG 523
Db 760 GAAGAAAGGAGGATAAAGTGATCAAGAGAT-GCAGCTCAGATCCGTGGAGTGGAGCAG 702
Qy 524 AGCCGACAGGAGTGGTTCGGTCTGTCTTAGAG----- 556
Db 701 AGCCGACAGGAGGT-GTTCGGTCTGTCTTAGAGGTTGTTCCTCCGAGGATCCAGACCA 643
Qy 557 CCTCAGGCGAGTGCAGACCCAGAGAGGGCTCTTCAGCACCTAGAGCTGGAAGGGATG 616
Db 642 CCTCAGGCGGT-CCAGACCCAGAGA-GGCTCTTCAGCACCTAGAGCTGGAAGGGATG 585
Qy 617 AACGCCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAATTTGGACCTGCCACAGCTCCA 676

Db 584 AACGCCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAATTTGGACCTGCCACAGCTCCA 525
Qy 677 GAGCTTGACTCGATGGAGACAGGACCACTCTCTGACATTCATTGGCCATCAGGATATACCA 736
Db 524 GAGCTTGACTCGATGGAGACAGGACCACTCTCTGACATTCATTGGCCATCAGGATATACCA 465
Qy 737 GGAGTTGGTAAACATCCACTCAGGTGCCACACTCCCTGGATGATCCAGATGAAGAATAC 796
Db 464 GGAGTTGGTAAACATCCACTCAGGTGCCACACTCCCTGGATGATCCAGATGAAGAATAC 405
Qy 797 ATTCTCTGGACCAAGAAATAGGACCACTCTATGAAGAAATTTCTTAAGAAAGGAAAA 856
Db 404 ATTCTCTGGACCAAGAAATAGGACCACTCTATGAAGAAATTTCTTAAGAAAGGAAAA 345
Qy 857 CAGAAGTTGAAAAAATCCCTCCAGACCGAGTGTGGGCAACTTTTGATCAGAGCTCCAGG 916
Db 344 CAGAAGTTGAAAAAATCCCTCCAGACCGAGTGTGGGCAACTTTTGATCAGAGCTCCAGG 285
Qy 917 ACCAGTGAGGCTGGCTGCCCTCTTTTGGGCGCGTCTGGAATTAATGGACCGCGTGGCA 976
Db 284 ACCAGTGAGGCTGGCTGCCCTCTTTT-GGCGCGTCTGGAATTAATGGACCGCGTGGCA 226
Qy 977 GTCCAGACATCAACTCCAAAATGAAGTGCAGCAATGAAGAAGCAGTCAATACAGAA 1036
Db 225 GTCCAGACATCAA-TTCAAAATGAAGTGCAGCAATGAAGAAGCAGTCAATACAGAA 167
Qy 1037 AAAGCTAATCATGCTCTTACCAACTACGAGTGTGGGCAACTTTTGATCAGAGCTCCAGG 1096
Db 166 AAAGCTAATCATGCTCTTACCAACTACGAGTGTGGGCAACTTTTGATCAGAGCTCCAGG 109
Qy 1097 CTATTATACCTTCCACCAAAATCTTTTATCATCTTTTCTTTAGGAAACAGACATATCTCA 1156
Db 108 CTATTATACCTTCCA-CCAAATCTTTTATCATCTTTTCTTTAGGAAACAGACATATCTCA 50
Qy 1157 TCGATTTGATTTAATAAGTTTATTTTC 1186
Db 49 TTCATTTGATTTAATAAGTTTATTTTC 20

RESULT 9
BI414452
LOCUS
DEFINITION
602986067F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5142137 5',
mRNA sequence.
ACCESSION
BI414452
VERSION
BI414452.1 GI:15175388
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 878)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: csapbe-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1350 row: e column: 18
High quality sequence start: 4
High quality sequence stop: 877.
FEATURES
source
Location/Qualifiers
1..878
/organism="Mus musculus"
/mol_type="mRNA"

QY 440 ATGGTGAAGGTTTGGATTCTTATCAAGAAAGGAGGATAAGTCAAGGAGATGGCA 499
 |||||
 Db 396 ATGGTGAAGGTTTGGATTCTTATCAAGAAAGGAGGATAAGTCAAGGAGATGGCA 455
 |||||
 QY 500 GCTCAGATCCGTGAGGTGAGGAGAGCCAGACAGGAGGTGGTTGGTCTGTCTTAGAGGCT 559
 |||||
 Db 456 GCTCAGATCCGTGAGGTGAGGAGAGCCAGACAGGAGGTGGTTGGTCTGTCTTAGAGGCT 515
 |||||
 QY 560 CAGGAGTGCAGAGCCAGAGAGAGGCTCTTTCAGCACTTAGAAGCTGAAAGGGATGAAC 619
 |||||
 Db 516 CAGGAGTGCAGAGCCAGAGAGAGGCTCTTTCAGCACTTAGAAGCTGAAAGGGATGAAC 575
 |||||
 QY 620 AGCCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAA--TTTGGACCTGCCACAGCTCCAG 677
 |||||
 Db 576 AGCCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGGACCTGCCACAGCTCCAG 635
 |||||
 QY 678 AGCTT 682
 |||||
 Db 636 AGCTT 640

RESULT 11
 BU608980/c
 LOCUS
 DEFINITION UI-CF-FNO-aer-m-01-0-UI.s1 743 bp mRNA linear EST 21-FEB-2003
 UI-CF-FNO-aer-m-01-0-UI 3', mRNA sequence.
 ACCESSION BU608980
 VERSION
 KEYWORDS
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 743)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: McCray, PB
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.reagen.com) or from Open Biosystems
 (www.openbiosystems.com).
 The following repetitive elements were found in this cDNA
 sequence: 1-47, >AT rich#Low_complexity
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source
 1. 743
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-CF-FNO-aer-m-01-0-UI"
 /tissue_type="Human Lung Epithelial cells"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-CF-FNO"
 /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
 modified polylinker; Site1: EcoR I; Site2: Not I;
 UI-CF-FNO is a subtracted cDNA library derived from two
 normalized Human lung epithelial cell libraries (BN1 and
 DU1) The library was subtracted according to according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. For additional information, contact:
 bent-soares@uiowa.edu
 TAG_SEQ=CTGCTCAGGT"
 TAG_L1B=UI-CF-FNO
 6hr to LPS 24h
 TAG_L1B=UI-CF-FNO
 TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match 44.6%; Score 558.4; DB 5; Length 743;
 Best Local Similarity 98.1%; Pred. No. 1.3e-141;
 Matches 618; Conservative 0; Mismatches 6; Indels 6; Gaps 5;
 QY 557 CCTCAGGAGTGCAGACCCAGAGAGGCTCTTTCAGCACTTAGAAGCTGAAAGGGATG 616
 |||||
 Db 649 CCTCAGGAGTGCAGACCCAGAGAGGG-TCTTTCAGCACCTTAGAAGCTGAAAGGGATG 591
 |||||
 QY 617 AACAGCCAAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGGACCTGCCACAGCTCCA 676
 |||||
 Db 590 AACAGCCAAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGGACCTGCCACAGCTCCA 531
 |||||
 QY 677 GAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGGCCATCAGGATATACCA 736
 |||||
 Db 530 GAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGGCCATCAGGATATACCA 471
 |||||
 QY 737 GGAGTTGGTAACTCCATCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGAATAC 796
 |||||
 Db 470 GGAGTTGGTAACTCCATCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGAATAC 411
 |||||
 QY 797 ATTGCTGGGACCAAGAAATAGGACCATCTTATGAAGAAATTTCTTAAAGAAAGGAAAA 856
 |||||
 Db 410 ATTGCTGGGACCAAGAAATAGGACCATCTTATGAAGAAATTTCTTAAAGAAAGGAAAA 351
 |||||
 QY 857 CAGAAGTTGAAAAAATCTCCGCCAGACGAGTGGGGCCAACTTTTGATCAGAGCTCCAGG 916
 |||||
 Db 350 CAGAAGTTGAAAAAATCTCCGCCAGACGAGTGGGGCCAACTTTTGATCAGAGCTCCAGG 291
 |||||
 QY 917 ACCAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGTCTGGAAATATGACGCGCTGGCA 976
 |||||
 Db 290 ACCAGTGCAGGCTGGCTGCCCTCTTTT-GGGCGGCTCTGGAAATATGACGCGCTGGCA 232
 |||||
 QY 977 GTCCAGACATCACTCCAAAATGAGCTGAGCAATGAAGAAGCAGTCAATACAGAAA 1036
 |||||
 Db 231 GTCCAGACATCAA-TTCAAAAATGAGCTGAGCAATGAAGAAGCAGTCAATACAGAAA 173
 |||||
 QY 1037 AAAGCTAATCATGCTCTCTACCAACTACCTAGGCTTAAAGCCAAAGTCAACCAACCC 1096
 |||||
 Db 172 AAAGCTAATCATGCTCTCTACCAACTACCTAGGCTTAAAGC--AAAGTCAACCAACCC 115
 |||||
 QY 1097 CTATTATACCTTCCACCCAAATTTCTTATCATTTCTTTTCTTAGGAAACAGACATCTCA 1156
 |||||
 Db 114 CTATTATACCTTCCA-CCAAATTTCTTATCATTTCTTTCTTAGGAAACAGACATCTCA 56
 |||||
 QY 1157 TTCATTTGATTTAATAAAGTTTATTTTTC 1186
 |||||
 Db 55 TTCATTTGATTTAATAAAGTTTATTTTTC 26
 |||||

RESULT 12

BU608895/c
 LOCUS
 DEFINITION UI-CF-FNO-aer-m-01-0-UI.s1 747 bp mRNA linear EST 21-FEB-2003
 UI-CF-FNO-aer-m-01-0-UI 3', mRNA sequence.
 ACCESSION BU608895
 VERSION
 KEYWORDS EST.
 SOURCE BU608895.1 GI:23275110
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 747)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL MEDLINE PUBMED COMMENT	Genome Res. 6 (9), 791-806 (1996) 97044477 8889548 Contact: McCray, PB McCray Lab University of Iowa 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA Tel: 319 356 4866 Fax: 319 356 7171 Email: paul-mccray@uiowa.edu Tissue Procurement: Dr. M. J. Welsh, University of Iowa CDNA Library prepared by: Dr. M. Bento Soares, University of Iowa CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com). The following repetitive elements were found in this cDNA sequence: 1-47, >AT rich#Low_complexity Seq primer: M13 FORWARD POLYA=Yes.	Db	290		ACCACTGAGGCTGGCTCCCTCTTTT-GGCCGGCTCTGGAATATGAGCGCGCTGGCA	232
		Qy	977		GTCCAGACATCAACTCCAAAACTGAAGCTGCAGCAATGAAGAAGCACTCATACAGAAA	1036
		Db	231		GTCCAGACATCAA-TTCAAACTGAAGCTGCAGCAATGAAGAAGCACTCATACAGAAA	173
		Qy	1037		AAAGTAAATCATGCTCTCTACCACTACCATGAGGCTAAAAGCAAGTCAACCAACCC	1096
		Db	172		AAAGTAAATCATGCTCTCTACCACTACCATGAGGCTAAAAGC--AAAGTCAACCAACCC	115
		Qy	1097		CTATTATACCTTCCACCCAAAATTTCTTATCATCTTCTTCTTAGGAAACAGACATCTCA	1156
		Db	114		CTATTATACCTTCCCA-CCAAATTTCTTATCATCTTCTTCTTAGGAAACAGACATCTCA	56
		Qy	1157		TTCAATTTCATTTAATAAAGTTTATTTTTC 1186	
		Db	55		TTCAATTTCATTTAATAAAGTTTATTTTTC 26	
FEATURES	Location/Qualifiers					
source	1. .747	RESULT 13				
	/organism="Homo sapiens"	BM684305/c				
	/mol_type="mRNA"	LOCUS				
	/db_xref="taxon:9606"	DEFINITION				
	/clone="UI-CF-FNO-aeg-1-18-0-UI"	ACCESSION				
	/tissue_type="Human Lung Epithelial cells"	VERSION				
	/lab_host="DH10B (Life Technologies) (T1 phage resistant)"	KEYWORDS				
	/clone_lib="UI-CF-FNO"	SOURCE				
	/note="Organ: Lung; Vector: pTT3-Pac (Pharmacia) with a	ORGANISM				
	modified polylinker; Site 1: EcoR I; Site 2: Not I;					
	UI-CF-FNO is a subtracted cDNA library derived from two	REFERENCE				
	normalized Human lung epithelial cell libraries (EN1 and	AUTHORS				
	DUI) The library was subtracted according to according to	TITLE				
	Bonaldo, Lennon and Soares, Genome Research, 6:791-806,	JOURNAL				
	1996. For additional information, contact:	MEDLINE				
	bento-soares@uiowa.edu	PUBMED				
	TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS	COMMENT				
	6hr_to LPS 24h					
	TAG_LIB=UI-CF-FNO					
	TAG_SEQ=CTGCTCAGGT"					
ORIGIN						
	Query Match 44.6%; Score 558.4; DB 5; Length 747;					
	Best Local Similarity 98.1%; Pred. No. 1.3e-141;					
	Matches 618; Conservative 0; Mismatches 6; Indels 6; Gaps 5;					
Qy	557		CCTCAGGCAAGTCCAGACCCAGAGAGGGCTCTTTCAGACCTCTAGAGCTGGAAGGGATG	616		
Db	649		CCTCAGGCAAGTCCAGACCCAGAGAGGG-TCTTCAGACCTCTAGAGCTGGAAGGGATG	591		
Qy	617		ACAGCCAAAGTAGTTCAGCTTACAGAGCCCTCAAAATTTGGACCTGCCACCAAGCTCCA	676		
Db	590		ACAGCCAAAGTAGTTCAGCTTACAGAGCCCTCAAAATTTGGACCTGCCACCAAGCTCCA	531		
Qy	677		GAGCTTGATCGATGGAGACAGACCATCTCTGACATTCATGGCCATCAGGATATACCA	736		
Db	530		GAGCTTGATCGATGGAGACAGACCATCTCTGACATTCATGGCCATCAGGATATACCA	471		
Qy	737		GGAGTTGGTAAATCAATCACTCAGTGCACACCTCCCTGGATGATCCAAAGTGAAGAATAC	796		
Db	470		GGAGTTGGTAAATCAATCACTCAGTGCACACCTCCCTGGATGATCCAAAGTGAAGAATAC	411		
Qy	797		ATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAGAAAGGAAAAA	856		
Db	410		ATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAGAAAGGAAAAA	351		
Qy	857		CAGAAGTTGAAAAAATCTCCCAAGCCAGAGTTGGGGCCAACTTTCATCAGCTCAGG	916		
Db	350		CAGAAGTTGAAAAAATCTCCCAAGCCAGAGTTGGGGCCAACTTTCATCAGCTCAGG	291		
Qy	917		ACCAGTGCAGGCTGGCTGGCTCTTTTGGGCGCGCTCTGGAATATGAGCGCGCTCGCA	976		

with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI). TAG_LIB=UI-E-EJ1 TAG_SEQ=GTCC"

ORIGIN

Query Match 43.8%; Score 548.6; DB 4; Length 656;
Best Local Similarity 96.8%; Pred. No. 6e-139;
Matches 613; Conservative 0; Mismatches 14; Indels 6; Gaps 5;
QY 554 GAGCTCAGGAGTCCAGACCCAGAGAGGGCTCTTCAGCACCTAGAGCTGGAAGGG 613
DB 552 GAGAGCTCAGGATGCCAGACCCAGAGA-GCTCTTCAGCACCTAGAGCTGGAAGGG 594
QY 614 ATGAACAGCCAAAGTAGCTTCCAGCTTACAGCAGGCCCTCAAAATTTGGACCTGCCACAGCT 673
DB 593 ATGAACAGCCAAAGTAGCTTCCAGCTTACAGCAGGCCCTCAAAATTTGGACCTGCCACAGCT 534
QY 674 CCAGAGCTTGAGTGAGAGAGAGGACCATCTCTGCATCTTTCAGCATTCATTGGCCATCAGGATATA 733
DB 533 CCAGAGCTTGAGTGAGAGAGAGGACCATCTCTGCATTCATTGGCCATCAGGATATA 474
QY 734 CCAGAGTTGTATACATCCATCAGGTGCCACACTCCCTGGATGATCCAGATGAAGAA 793
DB 473 CCAGAGTTGTATACATCCATCAGGTGCCACACTCCCTGGATGATCCAGATGAAGAA 414
QY 794 TACATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAGAAAGGAA 853
DB 413 TACATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAGAAAGGAA 354
QY 854 AAACAGAGTTGAAAAAATCCCTCCAGAGCCGAGTGGGGCCAACTTTGATCAAGCTCC 913
DB 353 AAACAGAGTTGAAAAAATCCCTCCAGAGCCGAGTGGGGCCAACTTTGATCAAGCTCC 294
QY 914 AGGACCAAGTCCAGCTGCTGCTCTTTTGGGGCGGCTCTGGAATATGACGCCCGCTG 973
DB 293 AGGACCAAGTCCAGCTGCTGCTCTTTTGGGGCGGCTCTGGAATATGACGCCCGCTG 235
QY 974 GCAGTCCAGACATCAACTCCAAAATCTGAAGCTGCAGCAATGAAGAAGCAGTTCATACAG 1033
DB 234 GCAGTCCAGACATCAA-TTCAAACTGAGCTGAGCAATGAAGAAGCAGTTCATACAG 176
QY 1034 AAAAAAGCTAATATGCTCTTACCAACTACCATGAGCTTAAAGCCAAAGTCAACAAA 1093
DB 175 AAAAAAGCTAATATGCTCTTACCAACTACCATGAGGCTAAAGC-AAAGTCAACAAA 118
QY 1094 CCCCTATATACCTTCCACCAATCTTTATCATCTCTTCTTAGGAACACACATAC 1153
DB 117 CCCCTATATACCTTCCCA-CCAAAATCTTTATCATCTCTTCTTAGGAACACACATAC 59
QY 1154 TCATTTCATTGATTAATAAAGTTTATTTTTC 1186
DB 58 TCATTTCATTGATTAATAAAGTTTATTTTTC 26

RESULT 14

BM929870 596 bp mRNA linear EST 13-MAR-2002
LOCUS UI-E-EJ1-aj1-1-23-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone
DEFINITION UI-E-EJ1-aj1-1-23-0-UI 5', mRNA sequence.
ACCESSION BM929870
VERSION BM929870.1 GI:19389043

KEYWORDS
SOURCE
ORGANISM

EST.
Homo sapiens (human)

Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 596)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene
discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

8889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 REVERSE.

Location/Qualifiers

source

1..596

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-E-EJ1-aj1-1-23-0-UI"

/tissue_type="fetal eyes, lens, eye anterior segment,

optic nerve, retina, Retina Foveal and Macular, RPE and

Choroid"

/dev_stage="fetal and adult"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-E-EJ1"

/notes="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site_1: EcoR I; Site_2: Not I;

UI-E-EJ1 is a subtracted cDNA library constructed

according to Bonaldo, Lennon and Soares, Genome Research,

6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dT primer containing a Not I site. Double

stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pT7T3-Pac

vector. The oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The

sequence tags for this library are: fetal eyes,

AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

AATGCCGAT; optic nerve, CCATTAGTG; retina, CCGCG; Retina

Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This

library was created for the program, Gene Discovery in the

Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Query Match 43.2%; Score 541.4; DB 5; Length 596;
Best Local Similarity 98.2%; Pred. No. 5.5e-137;
Matches 590; Conservative 0; Mismatches 6; Indels 5; Gaps 4;
QY 553 AGAGCCTCAGGAGTCCAGACCCAGAGAGGGCTCTTCAGCACCTAGAGCTGGAAGG 612
DB 1 AGAGCCTCAGGAGTCCAGACCCAGAGAGGGCTCTTCAGCACCTAGAGCTGGAAGG 60
QY 613 GATGAACAGCAAGTAGTCTCCAGCTTACAGCAGCCCTCAAAATTTGGACCTGCCACGAG 672
DB 61 GATGAACAGCAAGTAGTCTCCAGCTTACAGCAGCCCTCAAAATTTGGACCTGCCACGAG 120
QY 673 TCAGAGCTTGAGTGAGAGAGGACCATCTCTGACATTCATTGGCCATCAGGATAT 732
DB 121 TCAGAGCTTGAGTGAGAGAGGACCATCTCTGACATTCATTGGCCATCAGGATAT 180

QY 733 ACCAGAGTGTGTAACATCCACTCAGGTGGCCACACCTCCCTGGATGATCCAGATGAAGA 792
Db |||||||
181 ACCAGAGTGTGTAACATCCACTCAGGTGGCCACACCTCCCTGGATGATCCAGATGAAGA 240
QY 793 ATACATTCTGCGAACCAAGAAATAGGACCATCCTATGAGAAATTTCTTAAAGAAAGGA 852
Db |||||||
241 ATACATTCTGCGAACCAAGAAATAGGACCATCCTATGAGAAATTTCTTAAAGAAAGGA 300
QY 853 AAAACAGAGTTGAAAAAATCCCCCGCAGACCGAGTTGGGGCCAACTTTGATCACAGTTC 912
Db |||||||
301 AAAACAGAGTTGAAAAAATCCCCCGCAGACCGAGTTGGGGCCAACTTTGATCACAGTTC 360
QY 913 CAGGACCATGTCAGGCTGGCTGCCCTCTTTTGGGGCGGCTCTGGAATTAATGGACCGCGT 972
Db |||||||
361 CAGGACCATGTCAGGCTGGCTGCCCTCTTTT-GGGCGGCTCTGGAATTAATGGACCGCGT 419
QY 973 GCGAGTCCAGACATCACTCCAAACTGAAGCTGAGCAGCAATGAAGAGCAGTCAATACA 1032
Db |||||||
420 GCGAGTCCAGACATCAA-TTCAAAACTGAAGCTGAGCAGCAATGAAGAGCAGTCAATACA 478
QY 1033 GAAAAAGCTAATCATGCTCTCTACCACTACCATGAGCTAAAAAGCCAAAGTCAACCAA 1092
Db |||||||
479 GAAAAAGCTAATCATGCTCTCTACCACTACCATGAGCTAAAAGC--AAGTCAACAA 536
QY 1093 ACCCTATTATACCTTCCACCAAAATCTTTATCATTTGCTTTCTTAGGAAAACAGACATA 1152
Db |||||||
537 ACCCTATTATACCTTCCA-CCAAATTTCTTTATCATTTGCTTTCTTAGGAAAACAGACATA 595
QY 1153 C 1153
Db 596 C 596

RESULT 15
BI046266/c
LOCUS
DEFINITION
MR3-FN0209-060201-016-f06 FN0209 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BI046266
VERSION
BI046266.1 GI:14452888
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 668)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-FN0209-
060201-016-f06&t3=2001-02-06&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 668.
Location/Qualifiers
1..668
/organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="FN0209"
/note="Organ: prostate normal; Vector: puc18; Site: 1:
SmaI; Site 2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

ORIGIN

Query Match 42.4%; Score 531.8; DB 4; Length 668;
Best Local Similarity 91.6%; Pred. No. 2.4e-134;
Matches 513; Conservative 0; Mismatches 28; Indels 28; Gaps 4;
QY 481 AGTATCAAGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGCCGACAGGAGGTGT 540
Db |||||||
668 AGTATCAAGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGCCGACAGGAGGTGT 609
QY 541 TCGGTCTCTTTAGAGCCT-----CAGGCAGTGCAGACCCAG 578
Db |||||||
608 TCGTCTGTTTAGAGCTTGGTTTCCCTCGGAGGATCCAGACCAGTGCAGACCCAG 549
QY 579 AAGAGGGCTCTTCAGCACCTAGAAGCTGGAAGGGATGAACAGCCAAAGTAGCTTCCAGCT 638
Db |||||||
548 AAGAGGGCTCTTCAGCACCTAGAAGCTGGAAGGGATGAACAGCCAAAGTAGCTTCCAGCT 489
QY 639 TACAGAGCCCTCAAAATTTGGACCTGCCACAGCTCCAGAGCTTGACTGGATGGAGACAG 698
Db |||||||
488 TACAGAGCCCTCAAAATTTGGACCTGCCACAGCTCCAGAGCTTGACTGGATGGAGACAG 429
QY 699 GACCATCTCGACATTTTCATTTGGCCATCAGGATATACACAGGAGTTGGTAACTCCACTCAG 758
Db |||||||
428 GACCATCTCGACATTTTCATTTGGCCATCAGGATATACACAGGAGTTGGTAACTCCACTCAG 369
QY 759 GTGCCACACCTCCCTCGGATGATCCAAGATGAAGATACATTGCTGGGAACCAAGAAATAG 818
Db |||||||
368 GTGCCACACCTCCCTCGGATGATCCAAGATGAAGATACATTGCTGGGAACCAAGAAATAG 309
QY 819 GACCATCTCGACAGAAATTTCTTAAAGAAAAGGAAAAACAGAAAGTTGAAAAAATCCCCC 878
Db |||||||
308 GACCATCTCGACAGAAATTTCTTAAAGAAAAGGAAAAACAGAAAGTTGAAAAAATCCCCC 249
QY 879 CAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCAAGTGCAGCTGGCTGCCCT 938
Db |||||||
248 CAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCAAGTGCAGCTGGCTGCCCT 189
QY 939 CTTTTGGGCGCGCTCTGGAATAATGGACCGCGCTGGCAGTCCAGACATCAACTCCAAAAC 998
Db |||||||
188 CTTTT-GNCCGCGCTCTGGAATAATGGACCGCGCTGGCAGTCCAGACATCAA-TTCAAAAC 131
QY 999 TGAAGCTCGAGCAATGAAGAGCAGTCAATACAGAAAAAGCTAATCATGCTCTCTACC 1058
Db |||||||
130 TGAAGCTCGAGCAATGAAGAGCAGTCAATACAGAAAAAGCTAATCATGCTCTCTACC 71
QY 1059 AACTACCATGAGGCTAAAAGCCAAAGTCAACCAAAACCCCTATTATACCTTCCACCCAAAT 1118
Db |||||||
70 AACTACCATGAGGCTAAAAGCCAAAGTCAACCAAAACCCCTATTATACCTTCCACCTAGAC 15
QY 1119 TCCTTTATCA 1127
Db 14 TCCTTTCCA 6

Search completed: June 10, 2005, 20:27:43
Job time : 2650.16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:21:54 ; Search time 138.136 Seconds
(without alignments)
14842.245 Million cell updates/sec

Title: US-09-155-676B-4
Perfect score: 1253
Sequence: 1 cattggagtcacgcggtggc.....ggggggcgctaccacctttt 1253

Scoring table: IDENTITY NUC
Gapop 10'0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	278.4	22.2	816	4	US-09-949-016-3437
2	244.8	19.5	374	4	US-09-513-999C-22599
3	175.8	14.0	6623	4	US-09-949-016-16221
4	175.8	14.0	7031	4	US-09-949-016-15179
5	122.8	9.8	328	4	US-09-513-999C-34197
6	83.4	6.7	97	4	US-09-513-999C-15628
7	60.4	4.8	7218	1	US-08-232-463-14
8	46.8	3.7	2126	2	US-08-789-354-1
9	46.8	3.7	2126	3	US-09-110-937-1
10	46.8	3.7	2126	3	US-09-058-725B-1
11	46.8	3.7	2126	3	US-09-232-857-1
12	44.8	3.6	1236	2	US-08-933-115-1
13	44.8	3.6	1236	3	US-08-205-008-1
14	44.8	3.6	1236	4	US-09-206-115-1
15	43.8	3.5	1981	3	US-08-981-392-26
16	43.8	3.5	1981	4	US-09-908-322-26
17	43.2	3.4	1023	1	US-08-198-446B-1
18	43.2	3.4	1023	2	US-08-870-693-1
19	43.2	3.4	2150	1	US-08-198-446B-10
20	43.2	3.4	2150	2	US-08-870-693-10
21	43	3.4	852	1	US-08-075-533-21
22	43	3.4	852	2	US-08-948-176-21
23	43	3.4	852	5	PCT-US91-09160-21
24	42.4	3.4	637	4	US-09-387-286-25
25	41.8	3.3	470	4	US-10-101-464A-422
26	41.2	3.3	950	4	US-09-636-499-20
27	41	3.3	51259	3	US-08-781-891-209

c	28	41	3.3	51259	4	US-09-618-166-209	Sequence 209, App
	29	40.8	3.3	71278	4	US-09-949-016-11851	Sequence 11851, A
	30	40.8	3.3	71278	4	US-09-949-016-17563	Sequence 17563, A
c	31	40.8	3.3	168174	4	US-10-071-411A-63	Sequence 63, Appl
c	32	40.8	3.3	168273	4	US-10-071-411A-2	Sequence 2, Appl
	33	40.6	3.2	145	4	US-09-822-250A-6	Sequence 6, Appl
	34	40.6	3.2	145	4	US-10-034-350A-6	Sequence 6, Appl
	35	40.6	3.2	148	4	US-09-822-250A-7	Sequence 7, Appl
	36	40.6	3.2	148	4	US-10-034-350A-7	Sequence 7, Appl
	37	40.6	3.2	149	4	US-09-822-250A-8	Sequence 8, Appl
	38	40.6	3.2	149	4	US-10-034-350A-8	Sequence 8, Appl
	39	40.6	3.2	150	4	US-09-822-250A-9	Sequence 9, Appl
	40	40.6	3.2	150	4	US-10-034-350A-9	Sequence 9, Appl
	41	40.4	3.2	3751	4	US-09-140-378A-1	Sequence 1, Appl
	42	40.4	3.2	3803	4	US-09-023-655-1272	Sequence 1272, Ap
	43	40.4	3.2	3943	4	US-09-949-016-2495	Sequence 2495, Ap
	44	40.4	3.2	3943	4	US-09-949-016-2496	Sequence 2496, Ap
	45	40.4	3.2	3943	4	US-09-949-016-2497	Sequence 2497, Ap

ALIGNMENTS

RESULT 1

US-09-949-016-3437
; Sequence 3437, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE OF INVENTION: CLO01307
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3437
; LENGTH: 816
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3437

Query Match	22.2%	Score 278.4;	DB 4;	Length 816;
Best Local Similarity	96.7%	Pred. No. 1.1e-71;		
Matches	327;	Conservative	0;	Mismatches 6;
				Indels 5;
				Gaps 4;
Qy	849	AGGAAAAAGAGAGTTGAAAAAACTCCCCAGACCCAGTTGGGGCCCACTTTGATCACA	908	
Db	1	AGGAAAAAGAGAGTTGAAAAAACTCCCCAGACCCAGTTGGGGCCCACTTTGATCACA	60	
Qy	909	GCTCCAGACCCAGTGCAGGCTGGCTGCTCTTTTGGGCGCGCTCTGGAATAATGGAGCG	968	
Db	61	GCTCCAGACCCAGTGCAGGCTGGCTGCTCTTTTGGGCGCGCTCTGGAATAATGGAGCG	119	
Qy	969	CGCTGGCAGTCCAGACATCAACTCCAAAATGGAAGCTGCAGCAATGAAGAACGAGTCACA	1028	
Db	120	CGCTGGCAGTCCAGACATCAATTTCAAACTGGAAGCTGCAGCAATGAAGAACGAGTCACA	178	
Qy	1029	TACGAAAAAGCTAATCATGCTCTCTTACCACTACCATGAGGCTAAAGCAAGTCAA	1088	
Db	179	TACGAAAAAGCTAATCATGCTCTCTTACCACTACCATGAGGCTAAAGCAAGTCAA	236	
Qy	1089	CCAAACCCCTATTATACCTTCCACCCCAAAATCTTTTATCATTTGCTTTCTTAGGAAACAGA	1148	
Db	237	ACAAACCCCTATTATACCTTCCCA-CCAAAATCTTTTATCATTTGCTTTCTTAGGAAACAGA	295	
Qy	1149	CATACTCATTTGATTTTAATAAAGTTTTTATTTTC	1186	


```

; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ptzgpt-F1s
; US-08-232-463-14

Query Match 4.8%; Score 60.4; DB 1; Length 7218;
Best Local Similarity 5.3%; Pred.No. 9.9e-07;
Matches 22; Conservative 228; Mismatches 164; Indels 0; Gaps 0;

QY 243 TGTCGTCGGCTGTGAGGTGGGAAACACCTGAGCCATGGAAACCTGACGGTGTCTGTACG 302
DB 1442 TTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1383
QY 303 GGGGGCTGTGGAGCATCTGGCCAGCCGACGACACAAGAACCAACAATCTTGGT 362
DB 1382 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1323
QY 363 GGAGAGAACAAAGCTGAGGTCCAGATGAAGAGAGATTCTTGGTCACCTCCCGAGGATTG 422
DB 1322 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1263
QY 423 CGCGATTCAAGAAATCCATCGTGAAGAGTTGGATTCCTATGAAGAAAAGGAGATAAG 482
DB 1262 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1203
QY 483 TGATCAAGGAGATGGCAGCTCAGATCCGTGAGGTGGAGACGACCGACAGAGGTGTT 542
DB 1202 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1143
QY 543 GGTCTGTCTTAGAGCCTCAGGCGATCGCAGACCCAGAGAGGGCTCTTCAGCACCTAGAA 602
DB 1142 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1083
QY 603 GCTGGAAAGGATGAACAGCAAAGTAGCTTCAGCTTACAGCAGCCCTCAAAAT 656
DB 1082 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1029

```

```

US-09-110-937-1
Query Match      3.7%; Score 46.8; DB 3; Length 2126;
Best Local Similarity 96.0%; Pred. No. 0.0049;
Matches 48; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1204 ATCAAGCTTATCATACCGTCGACCTCGAGGGGGGGCGGTACCCACTTTT 1253
          |||||||
          53 ATCAAGCTTATCATACCGTCGACCTCGAGGGGGGGCGGTACCCAGTTT 4

RESULT 10
US-09-058-725B-1/c
; Sequence 1, Application US/09058725B
; Patent No. 6133420
; GENERAL INFORMATION:
; APPLICANT: Ames, Robert
; APPLICANT: Sarau, Henry
; APPLICANT: Foley, James
; APPLICANT: Chamber, Jon
; TITLE OF INVENTION: A Method of Finding Angonist
; TITLE OF INVENTION: and Antagonist To Human and Rat GPR 14
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
;

```

ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/058,725B
FILING DATE: April 10, 1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/789,354
FILING DATE: 27-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T.
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: GP50005-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-058-725B-1

Query Match 3.7%; Score 46.8; DB 3; Length 2126;
Best Local Similarity 96.0%; Pred. No. 0.0049;
Matches 48; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1204 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGCGTACCCACTTTT 1253
Db 53 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGCGTACCCACTTTT 4

RESULT 11
US-09-232-857-1/c
Sequence 1, Application US/09232857
Patent No. 6159700
GENERAL INFORMATION:
APPLICANT: DOUGLAS, STEPHEN
APPLICANT: WILLETTTE, ROBERT
APPLICANT: AIYAR, NAMBI
APPLICANT: ROMANIC, ANNE
APPLICANT: KHANDOUDI, NASSIRAH
APPLICANT: GOUT, BERNARD
APPLICANT: AL-BARAZANJI, KAMAL
APPLICANT: AMES, ROBERT S.
APPLICANT: FOLEY, JAMES J.
APPLICANT: SARAU, HENRY
APPLICANT: CHAMBERS, JON K.
APPLICANT: SHABON, USMAN
APPLICANT: BERGSMAN, DEREK
TITLE OF INVENTION: A METHOD OF FINDING AGONIST
AND ANTAGONIST TO HUMAN AND RAT GPR14
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ratner & Prestia
STREET: P.O. Box 980
CITY: Valley Forge
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/232,857
FILING DATE: 15-JAN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/789,354
FILING DATE: 27-JAN-1997
APPLICATION NUMBER: 60/074,075
FILING DATE: 09-FEB-1998
APPLICATION NUMBER: 09/058,725
FILING DATE: 10-APR-1998
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F.
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GP-50005-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0700
TELEX: 846169
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-232-857-1

Query Match 3.7%; Score 46.8; DB 3; Length 2126;
Best Local Similarity 96.0%; Pred. No. 0.0049;
Matches 48; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1204 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGCGTACCCACTTTT 1253
Db 53 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGCGTACCCACTTTT 4

RESULT 12
US-08-933-115-1/c
Sequence 1, Application US/08933115
Patent No. 5866369
GENERAL INFORMATION:
APPLICANT: Zalacain, Magdalena
APPLICANT: Brown, James R.
TITLE OF INVENTION: NOVEL AmpS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,115
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd Q
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: GM10099

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-933-115-1
Query Match 3.6%; Score 44.8; DB 2; Length 1236;
Best Local Similarity 95.8%; Pred. No. 0.014; 2; Indels 0; Gaps 0;
Matches 46; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1203 AATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGGTACCCCACT 1250
Db 48 AATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGGTACCCCAAT 1

RESULT 13
US-09-205-008-1/c
; Sequence 1, Application US/09205008
; Patent No. 6297039
; GENERAL INFORMATION:
; APPLICANT: Zalacain, Magdalena
; APPLICANT: Brown, James R.
; TITLE OF INVENTION: NOVEL Amps
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/205,008
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/933,115
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd O
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: GM10099
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-09-205-008-1
Query Match 3.6%; Score 44.8; DB 3; Length 1236;
Best Local Similarity 95.8%; Pred. No. 0.014;
Matches 46; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1203 AATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGGTACCCCACT 1250
Db 48 AATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGGTACCCCAAT 1

RESULT 14
US-09-206-115-1/c
; Sequence 1, Application US/09206115
; Patent No. 6558919
; GENERAL INFORMATION:
; APPLICANT: Zalacain, Magdalena
; APPLICANT: Brown, James R.
; TITLE OF INVENTION: NOVEL Amps
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/206,115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/933,115
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd O
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: GM10099
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-09-206-115-1
Query Match 3.6%; Score 44.8; DB 4; Length 1236;
Best Local Similarity 95.8%; Pred. No. 0.014;
Matches 46; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1203 AATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGGTACCCCACT 1250
Db 48 AATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGGTACCCCAAT 1

RESULT 15
US-08-981-392-26/c
; Sequence 26, Application US/08981392
; Patent No. 6262025
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos Manuel Pinto
; APPLICANT: Lewis, Julian Hart
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Gray, Grace
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES
; OF VERTEBRATE DELTA GENES AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
```



```
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,392
; FILING DATE: 22-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Antler, Adriane M.
; REGISTRATION NUMBER: 32,605
; REFERENCE/DOCKET NUMBER: 7326-038
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1981 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-981-392-26

Query Match      3.5%; Score 43.8; DB 3; Length 1981;
Best Local Similarity 95.7%; Pred. No. 0.036;
Matches 45; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1204 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGGTACCCACT 1250
    |||||||||||||||||||||||||||||||||||||||||||
Db 48 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGGTACCCCAAT 2
```

Search completed: June 10, 2005, 20:46:47
Job time : 140.386 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 14:23:14 ; Search time 476.179 Seconds
(without alignments)
16311.995 Million cell updates/sec

Title: US-09-155-676B-4
Perfect score: 1253
Sequence: 1 cattggagtcacgctggc.....ggggggcgctaccactttt 1253

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 309930249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	451.8	36.1	3948	9	US-09-915-543-16
2	451.8	36.1	3948	15	Sequence 16, Appl
3	451.8	36.1	3948	17	Sequence 16, Appl
4	444.4	35.5	500	17	US-10-664-859-16
5	444.4	35.5	500	18	US-10-242-535A-45728
6	400.8	32.0	2556	10	US-10-085-783A-45728
7	256.4	20.5	352	9	US-09-796-692-3344

8	256.4	20.5	352	14	US-10-040-862-3344	Sequence 3344, Ap
9	256.4	20.5	352	17	US-10-057-475B-3344	Sequence 3344, Ap
10	256.4	20.5	352	17	US-10-154-884B-3344	Sequence 3344, Ap
11	256.4	20.5	352	19	US-10-764-324-3344	Sequence 3344, Ap
12	219.6	17.5	970	13	US-10-027-632-120356	Sequence 120356,
13	219.6	17.5	970	17	US-10-027-632-120356	Sequence 120356,
c 14	205	16.4	567	21	US-10-956-157-3788	Sequence 3788, Ap
15	205	16.4	567	21	US-10-956-157-3788	Sequence 3788, Ap
16	186.4	14.9	254	17	US-10-342-535A-52797	Sequence 52797, A
17	186.4	14.9	254	18	US-10-085-783A-52797	Sequence 52797, A
18	104.8	8.4	258	17	US-10-131-827-8756	Sequence 8756, Ap
19	45.8	3.7	864	14	US-10-198-846-3416	Sequence 3416, Ap
20	45	3.6	647	21	US-10-643-775-1354	Sequence 1354, Ap
21	44.4	3.5	437	20	US-10-357-930-31854	Sequence 31854, A
c 22	44.2	3.5	1488	19	US-10-437-963-42386	Sequence 42386, A
c 23	43.8	3.5	1981	9	US-09-908-322-26	Sequence 26, Appl
c 24	43.8	3.5	1981	20	US-09-783-931-26	Sequence 26, Appl
c 25	43.8	3.5	1981	20	US-10-877-563-16	Sequence 16, Appl
26	43.4	3.5	1896	19	US-10-437-963-90219	Sequence 90219, A
27	43.2	3.4	468	20	US-10-357-930-31864	Sequence 31864, A
28	43.2	3.4	468	20	US-10-357-930-40827	Sequence 40827, A
29	43.2	3.4	478	20	US-10-357-930-31862	Sequence 31862, A
30	43.2	3.4	495	20	US-10-357-930-31845	Sequence 31845, A
31	43.2	3.4	495	20	US-10-357-930-40811	Sequence 40811, A
32	43.2	3.4	575	14	US-10-198-846-4142	Sequence 4142, Ap
c 33	43.2	3.4	805	14	US-10-198-846-7497	Sequence 7497, Ap
c 34	43	3.4	810	21	US-10-643-775-325	Sequence 325, App
35	43	3.4	810	21	US-10-643-775-668	Sequence 668, Appl
c 36	42.4	3.4	637	16	US-10-300-341-25	Sequence 25, Appl
37	42.2	3.4	860	14	US-10-198-846-4856	Sequence 4856, Ap
38	41.8	3.3	470	14	US-10-101-464A-422	Sequence 422, App
39	41.8	3.3	470	21	US-10-864-252-422	Sequence 422, App
40	41.6	3.3	1664	19	US-10-437-963-100273	Sequence 100273,
41	41.4	3.3	526	21	US-10-643-775-554	Sequence 554, App
42	41.2	3.3	783	21	US-10-643-775-1030	Sequence 1030, Ap
43	41.2	3.3	824	14	US-10-198-846-9085	Sequence 9085, Ap
c 44	41.2	3.3	950	15	US-10-229-358-20	Sequence 20, Appl
c 45	41	3.3	416	9	US-09-924-035A-641	Sequence 641, App

ALIGNMENTS

RESULT 1

US-09-915-543-16
; Sequence 16, Application US/0915543
; Publication No. US20020086986A1
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich
; APPLICANT: FROESCH, Barbara
; APPLICANT: KRAMPS, Thomas
; APPLICANT: PETER, Oliver
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WIRELESS SIGNALING PATHWAY
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
; FILE REFERENCE: Q60361
; CURRENT APPLICATION NUMBER: US/09/915,543
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,502
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 3948
; TYPE: DNA
; ORGANISM: Human lgs-1
; US-09-915-543-16

Query Match 36.1%; Score 451.8; DB 9; Length 3948;

Best Local Similarity 99.2%; Pred. No. 6.8e-126;

Matches 475; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 81 CCTTCTTCTGTGTGCG-GGCACGTTTACAGCCGCAAGCACCAGCGCAGCTGAAGGAG 139

Qy 380 GTCAGATGAAGAGAAAGTTCTGGTCACTCCCGAGGATTATGCGGANTCAAGAAATCC 439
Db |||||
3379 GTCCAGATGAAGAGAAAGTTCTGGTCACTCCCGAGGATTATGCGGANTCAAGAAATCC 3438
Qy 440 ATGGTGAAGGTTGGATTCTCTATCAAGAAAGAGGATAAAGTGCATCAAGGAGATGCA 499
Db |||||
3439 ATGGTGAAGGTTGGATTCTCTATCAAGAAAGAGGATAAAGTGCATCAAGGAGATGCA 3498
Qy 500 GCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGAGAGTGGTTCGCTCTGTCTTAGAGCC 558
Db |||||
3499 GCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGAGAGTGGTTCGCTCTGTCTTAGAGAC 3557

RESULT 4

US-10-242-535A-45728
; Sequence 45728, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 2001-07-13
; PRIOR FILING DATE: 2001-07-13
; PRIOR FILING DATE: 2001-03-12
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45728
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-45728

Query Match 35.5%; Score 444.4; DB 17; Length 500;
Best Local Similarity 97.8%; Pred. No. 3.7e-124; Indels 5; Gaps 4;
Matches 493; Conservative 0; Mismatches 6

Qy 683 GACTGGATGGAGACAGAGACCATCTCTGACATTCATTGGCCATCAGGATATACAGGAGTT 742
Db 1 GACTGGATGGAGACAGAGACCATCTCTGACATTCATTGGCCATCAGGATATACAGGAGTT 60
Qy 743 GGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAAAGTGAAGATACATTGCT 802
Db 61 GGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAAAGTGAAGATACATTGCT 120
Qy 803 GGGAAACCAAGAAATAGGACCATCTCTATGAAGAAATTTCTTAAAGAAAGGAAAAACAGAG 862
Db 121 GGGAAACCAAGAAATAGGACCATCTCTATGAAGAAATTTCTTAAAGAAAGGAAAAACAGAG 180
Qy 863 TTGAAAAAATCCCTCCCGACAGCCGAGTTGGGCGCCAACTTTTGATCAAGCTCCAGACCACT 922
Db 181 TTGAAAAAATCTCCCGACAGCCGAGTTGGGCGCCAACTTTTGATCAAGCTCCAGACCACT 240
Qy 923 GCAGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGAATAATGGACCGCGCTGGCAGTCCAG 982
Db 241 GCAGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGAATAATGGACCGCGCTGGCAGTCCAG 299
Qy 983 ACATCAACTCCAAATCGAAGCTGAGCAATGAAGAGCAGTCAATACAGAAAAAGCT 1042
Db 359 AATCATGCTCTCTACCAACTACCATGAGGCTTAAAGCCAAAGTCAACCAACCCCTATTA 1102
Qy 1103 TACCTTCCACCCAAATCTTTTATCATTTGCTTTTCTTAGGAAACAGACATCTCATTTCA 1162
Db 417 TACCTTCCA-CCAAATCTTTTATCATTTGCTTTTCTTAGGAAACAGACATCTCATTTCA 475
Qy 1163 TGATTTAATAAGTTTATTTTTC 1186
Db 476 TGATTTAATAAGTTTATTTTTC 499

Qy 1103 TACCTTCCACCCAAATCTTTTATCATTTGCTTTTCTTAGGAAACAGACATCTCATTTCA 1162
Db 417 TACCTTCCA-CCAAATCTTTTATCATTTGCTTTTCTTAGGAAACAGACATCTCATTTCA 475
Qy 1163 TGATTTAATAAGTTTATTTTTC 1186
Db 476 TGATTTAATAAGTTTATTTTTC 499

RESULT 5

US-10-085-783A-45728
; Sequence 45728, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 2001-07-13
; PRIOR FILING DATE: 2001-07-13
; PRIOR FILING DATE: 2001-03-12
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45728
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-45728

Query Match 35.5%; Score 444.4; DB 18; Length 500;
Best Local Similarity 97.8%; Pred. No. 3.7e-124; Indels 5; Gaps 4;
Matches 493; Conservative 0; Mismatches 6

Qy 683 GACTGGATGGAGACAGAGACCATCTCTGACATTCATTGGCCATCAGGATATACAGGAGTT 742
Db 1 GACTGGATGGAGACAGAGACCATCTCTGACATTCATTGGCCATCAGGATATACAGGAGTT 60
Qy 743 GGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAAAGTGAAGATACATTGCT 802
Db 61 GGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAAAGTGAAGATACATTGCT 120
Qy 803 GGGAAACCAAGAAATAGGACCATCTCTATGAAGAAATTTCTTAAAGAAAGGAAAAACAGAG 862
Db 121 GGGAAACCAAGAAATAGGACCATCTCTATGAAGAAATTTCTTAAAGAAAGGAAAAACAGAG 180
Qy 863 TTGAAAAAATCTCCCGACAGCCGAGTTGGGCGCCAACTTTTGATCAAGCTCCAGACCACT 922
Db 181 TTGAAAAAATCTCCCGACAGCCGAGTTGGGCGCCAACTTTTGATCAAGCTCCAGACCACT 240
Qy 923 GCAGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGAATAATGGACCGCGCTGGCAGTCCAG 982
Db 241 GCAGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGAATAATGGACCGCGCTGGCAGTCCAG 299
Qy 983 ACATCAACTCCAAATCGAAGCTGAGCAATGAAGAGCAGTCAATACAGAAAAAGCT 1042
Db 300 ACATCAA-TTCAAAACTGAAGCTGACCAATGAAGAGCAGTCAATACAGAAAAAGCT 358
Qy 1043 AATCATGCTCTCTACCAACTACCATGAGGCTTAAAGCCAAAGTCAACCAACCCCTATTA 1102
Db 359 AATCATGCTCTCTACCAACTACCATGAGGCTTAAAGCCAAAGTCAACCAACCCCTATTA 416
Qy 1103 TACCTTCCACCCAAATCTTTTATCATTTGCTTTTCTTAGGAAACAGACATCTCATTTCA 1162
Db 417 TACCTTCCA-CCAAATCTTTTATCATTTGCTTTTCTTAGGAAACAGACATCTCATTTCA 475
Qy 1163 TGATTTAATAAGTTTATTTTTC 1186
Db 476 TGATTTAATAAGTTTATTTTTC 499

RESULT 6

US-09-746-783-205
; Sequence 205, Application US/09746783
; Publication No. US2003004935A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.
Howes, Steven H.
Fechtel, Kim

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

ENCODING THEM

NUMBER OF SEQUENCES: 231
CORRESPONDENCE ADDRESS:
ADDRESSER: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/746,783
FILING DATE: 21-Dec-2000
CLASSIFICATION: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Milasincic, Debra J.
REGISTRATION NUMBER: 46,931
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214

INFORMATION FOR SEQ ID NO: 205:

SEQUENCE CHARACTERISTICS:
LENGTH: 2556 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 205:

US-09-746-783-205

Query Match 32.0%; Score 400.8; DB 10; Length 2556;
Best Local Similarity 97.4%; Pred. No. 1.8e-110;
Matches 450; Conservative 0; Mismatches 7; Indels 5; Gaps 4;

QY	725	CAGGATATACAGGAGTGGTAACTCCACTCAGGTGCCACACCTCCCTGGATGATCCAA	784
Db	2030	CAGGATATACAGGAGTGGTAACTCCACTCAGGTGCCACACCTCCCTGGATGATCCAA	2089
QY	785	GATGAAGAATACTGCTGGGAACCAAGAAATAGGACCATCTCTGAAGAATTTCTTAAA	844
Db	2090	GATGAAGAATACTGCTGGGAACCAAGAAATAGGACCATCTCTGAAGAATTTCTTAAA	2149
QY	845	GAAGAGGAAGAACAGAGTTGAAGAACTCCCGCCAGACCGAGTTGGGGCCAACTTGAT	904
Db	2150	GAAGAGGAAGAACAGAGTTGAAGAACTCCCGCCAGACCGAGTTGGGGCCAACTTGAT	2209
QY	905	CACAGCTCCAGGACCAAGTCCAGGCTGGCTCCCTCTTTTGGCCCGCTCTGGAATAATGG	964
Db	2210	CACAGCTCCAGGACCAAGTCCAGGCTGGCTCCCTCTTTT-GGCCGTGCTGGAATAATGG	2268
QY	965	ACGCCGTGGCGAGTCCAGACATCACTCCAAAACCTGAAGCTGCAGCAATGAAGACGAGT	1024

Db	2269	ACGCCGTGGCGAGTCCAGACATCAA-TTCAAAACTGAAGCTGCAGCAATGAAGACGAGT	2327
QY	1025	CACATACAGAAAAAGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCCAAAG	1084
Db	2328	CACATACAGAAAAAGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGC-AAA	2385
QY	1085	TCAACCAAAACCCCTATTATACCTTCCACCCTAAATTTCTTTATCATTTGTCTTTCTTAGGAAA	1144
Db	2386	GTCACAAACCCCTATTATACCTTCCA-CCAAATTTCTTTATCATTTGTCTTTCTTAGGAAA	2444
QY	1145	CAGACATACCTCATTCATTGTGATTTAATAAAGTTTTATTTTTC	1186
Db	2445	CAGACATACCTCATTCATTGTGATTTAATAAAGTTTTATTTTTC	2486

RESULT 7

US-09-796-692-3344
; Sequence 3344, Application US/09796692
; Publication No. US20020198362A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY

TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES

FILE REFERENCE: 2077.001200

CURRENT APPLICATION NUMBER: US/09/796,692

CURRENT FILING DATE: 2001-03-01

PRIOR APPLICATION NUMBER: 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: 60/222,903

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: 60/223,416

PRIOR FILING DATE: 2000-08-04

PRIOR APPLICATION NUMBER: 60/223,378

PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 9597

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 3344

LENGTH: 352

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: unsure

LOCATION: (92)

OTHER INFORMATION: n=A, T, C or G

NAME/KEY: unsure

LOCATION: (187)

OTHER INFORMATION: n=A, T, C or G

NAME/KEY: unsure

LOCATION: (235)

OTHER INFORMATION: n=A, T, C or G

NAME/KEY: unsure

LOCATION: (242)

OTHER INFORMATION: n=A, T, C or G

NAME/KEY: unsure

LOCATION: (318)


```
/ APPLICANT: Wang, Aijun
/ APPLICANT: Ordonez, Nadia
/ APPLICANT: Carter, Lauren
/ APPLICANT: McNeill, Patricia Dianne
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE OF INVENTION: Hematological Malignancies
/ FILE REFERENCE: 014058-014402US
/ CURRENT APPLICATION NUMBER: US/10/057,475B
/ CURRENT FILING DATE: 2002-01-22
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3344
/ LENGTH: 352
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(352)
/ OTHER INFORMATION: n = g, a, c or t
US-10-057-475B-3344
```

```
Query Match      20.5%; Score 256.4; DB 17; Length 352;
Best Local Similarity 97.2%; Pred. No. 5.1e-67;
Matches 312; Conservative 0; Mismatches 4; Indels 5; Gaps 5;
```

```
QY 324 CCAGCCCGAGCACAGAAAGCAACCAAAATTCGTGGGAGAACAAAGCTGAGGTCC 383
Db 1 CCAGCCCGAGCACAGAAAGCAACCAAAATTCGTGGGAGAACAAAGCTGAGGTCC 60

QY 384 AGATGAAAGAGAAGTTCTCGTCACTCCCGAGGATTATGCGCGATTCAAGAAATCCATGG 443
Db 61 AGATGAAAGAGAAGTTCTCGTCACTCCCGAGGATTATGCGCGATTCAAGAAATCCATGG 120

QY 444 TGAAGGTTTGATTCCTATGAAGAAAAGGAGGATAAAGTGATCAAGAGATGGCAGTCC 503
Db 121 TGAAGGTTTGATTCCTATGAAGAAAAGGAGGATAAAGTGATCAAGAGATGGCAGTCC 180

QY 504 AGATCC-GTGAGGTGGAGCAGAG-CCGACAGGAGTGGTTCGGTCTGCTTTAGA-CCCTC 560
Db 181 ACATCCNGTGGAGTGGAGCAGAGCCCGACAGGAGTGGTTCGGTCTGCTTTAGANGCCTC 240

QY 561 AGGCAGTGCCAGACCCAGAAAGGGCTCTTCAGCACCTAGAAAGC-TGGAAGGGATGAAC 619
Db 241 ANGCAAGTGCAGACCCAGAAAGGGCTCTTCAGCACCTAGAAAGC-TGGAAGGGATGAAC 300

QY 620 AG-CCAAGTAGTTCACGTT 639
Db 301 AGCCCAAGTAGTTCACNCTT 321
```

RESULT 10

```
US-10-154-884B-3344
/ Sequence 3344, Application US/10154884B
/ Publication No. US20040005561A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algarte, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE OF INVENTION: Hematological Malignancies
/ FILE REFERENCE: 014058-013521US
/ CURRENT APPLICATION NUMBER: US/10/154,884B
/ CURRENT FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 11290
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3344
/ LENGTH: 352
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(352)
/ OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-3344

Query Match      20.5%; Score 256.4; DB 17; Length 352;
Best Local Similarity 97.2%; Pred. No. 5.1e-67;
Matches 312; Conservative 0; Mismatches 4; Indels 5; Gaps 5;

QY 324 CCAGCCCGAGCACAGAAAGCAACCAAAATTCGTGGGAGAACAAAGCTGAGGTCC 383
Db 1 CCAGCCCGAGCACAGAAAGCAACCAAAATTCGTGGGAGAACAAAGCTGAGGTCC 60

QY 384 AGATGAAAGAGAAGTTCTCGTCACTCCCGAGGATTATGCGCGATTCAAGAAATCCATGG 443
Db 61 AGATGAAAGAGAAGTTCTCGTCACTCCCGAGGATTATGCGCGATTCAAGAAATCCATGG 120

QY 444 TGAAGGTTTGATTCCTATGAAGAAAAGGAGGATAAAGTGATCAAGAGATGGCAGTCC 503
Db 121 TGAAGGTTTGATTCCTATGAAGAAAAGGAGGATAAAGTGATCAAGAGATGGCAGTCC 180

QY 504 AGATCC-GTGAGGTGGAGCAGAG-CCGACAGGAGTGGTTCGGTCTGCTTTAGA-CCCTC 560
Db 181 ACATCCNGTGGAGTGGAGCAGAGCCCGACAGGAGTGGTTCGGTCTGCTTTAGANGCCTC 240

QY 561 AGGCAGTGCCAGACCCAGAAAGGGCTCTTCAGCACCTAGAAAGC-TGGAAGGGATGAAC 619
Db 241 ANGCAAGTGCAGACCCAGAAAGGGCTCTTCAGCACCTAGAAAGC-TGGAAGGGATGAAC 300

QY 620 AG-CCAAGTAGTTCACGTT 639
Db 301 AGCCCAAGTAGTTCACNCTT 321
```

Db 301 AGCCCAAGTAGTTCANCNTT 321

RESULT 11

US-10-764-324-3344

Sequence 3344, Application US/10764324

Publication No. US20040175739A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander

APPLICANT: Algate, Paul A.

APPLICANT: Mannion, Jane

APPLICANT: Retter, Marc

APPLICANT: Corixa Corporation

TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy

FILE REFERENCE: 014058-013520US

CURRENT APPLICATION NUMBER: US/10/764,324

CURRENT FILING DATE: 2004-01-23

PRIOR APPLICATION NUMBER: US/10/040,862

PRIOR FILING DATE: 2001-11-06

PRIOR APPLICATION NUMBER: US 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: US 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: US 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: US 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: US 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: US 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: US 60/218,950

PRIOR FILING DATE: 2000-07-14

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 10467

SOFTWARE: FastSEQ for Windows Version 3.0

SEQ ID NO 3344

LENGTH: 352

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: unsure

LOCATION: (92)

OTHER INFORMATION: n=A,T,C or G

FEATURE:

NAME/KEY: unsure

LOCATION: (187)

OTHER INFORMATION: n=A,T,C or G

FEATURE:

NAME/KEY: unsure

LOCATION: (235)

OTHER INFORMATION: n=A,T,C or G

FEATURE:

NAME/KEY: unsure

LOCATION: (242)

OTHER INFORMATION: n=A,T,C or G

FEATURE:

NAME/KEY: unsure

LOCATION: (318)

OTHER INFORMATION: n=A,T,C or G

FEATURE:

NAME/KEY: unsure

LOCATION: (329)

OTHER INFORMATION: n=A,T,C or G

US-10-764-324-3344

Query Match 20.5%; Score 256.4; DB 19; Length 352;

Best Local Similarity 97.2%; Pred. No. 5.1e-67;

Matches 312; Conservative 0; Mismatches 4; Indels 5; Gaps 5;

Qy 324 CCACCCAGAGCACAAGAAAGCAACCAAAATCTGTGGGAGAAACAAAGCTGAGGTCC 383

Db 1 CCAGCCAGAGCACAAGAAAGCAACCAAAATCTGTGTGGGAGAAACAAAGCTGAGGTCC 60

Qy 384 AGATGAAGAGAAAGTTTCTGTCTACCTCCCAAGGATTATGCCGATTCAAGAAATCCATGG 443

Db 61 AGATGAAGAGAAAGTTTCTGTCTACCTCCCAAGGATTATGCCGATTCAAGAAATCCATGG 120

Qy 444 TGAAGGTTTGGATTTCCTATGAAGAAAGGAGGATAAGTCAATCAAGGAGATGGCAGCTC 503

Db 121 TGAAGGTTTGGATTTCCTATGAAGAAAGGAGGATAAGTCAATCAAGGAGATGGCAGCTC 180

Qy 504 AGATCC-GTGAGGTGGAGCAGAG-CCGACAGAGAGGTGTTTCGGTCTGTCTTTAGA-GCCTC 560

Db 181 ACATCCNGTGGAGTGGAGCAGAGCCGACAGGAGTGGTTCGGTCTGTCTTTAGANGCCTC 240

Qy 561 AGCAGTGCACAGCCCAAGAGAGGGCTCTTCAGCACCTAGAGGC-TGAAAAGGGATGAAC 619

Db 241 ANGCAGTGCCAGACCCCAAGAGAGGGCTCTTCAGCACCTAGAGGC-TGAAAAGGGATGAAC 300

Qy 620 AG-CCAAAGTAGTTCACAGCTT 639

Db 301 AGCCCAAGTAGTTCANCNTT 321

RESULT 12

US-10-027-632-120356

Sequence 120356, Application US/10027632

Publication No. US20020198371A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 120356

LENGTH: 970

TYPE: DNA

ORGANISM: Human

US-10-027-632-120356

Query Match 17.5%; Score 219.6; DB 13; Length 970;

Best Local Similarity 81.8%; Pred. No. 1.5e-55;

Matches 332; Conservative 1; Mismatches 50; Indels 23; Gaps 6;

Qy 801 CTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAGAAAGGAAAAACAGA 860

Db 3 CTGGGAACCAAGAAACAGAGCCGCTCTATTAAGAAATTTATTAAGAAAGGAAAAACAGA 62

Qy 861 AGTTGAAAAAACTCCCCCAGAGCCGAGTGTGGGCCAACTTTTGATCACAGCTCCAGGACCA 920

Db 63 AGTTGAAAGGAACCTTTCCCAAACTGAGTTGGGGCCAACTTTAATCACAGCTCCAGGACCA 122

Qy 921 GTGCAGGCTGGCTGCCCTCTTTTGGGGCCGGTCTGGAATAATGACGCCGCTGCAGTCC 980

Db 123 GTCCAGGCTGGCTGCCCYC-|T|T|G|G|A|C|C|G|T|G|T|G|G|A|T|A|A|T|G|G|A|T|A|C|T|G|T|T|G|G|C|A|G|T|T|C 181
Qy 981 AGACATCACTCCAAACTGAAGCTGCAGCAATGAAGAAGCAGTCA-|-----CATAC 1031
Db 182 AGCGTCACT-CAGAACTGAAGCTGCAGCAATGAAGAAGCAGTCAATTA 240
Qy 1032 AGAAAAAGCTAATCATGC-|-----T|C|T|C|A|C|A|A|C|T|A|C|C|A|G|G|C|T|A|A|A|A|G|C|-----C 1080
Db 241 AAAAAATGTTAATCATTCCTGATGCTCTACCACTACCGTGAGGCTAAAGCAAGTC 300
Qy 1081 AAGTCAACCAACCCTATTATPACCTTCACCCAAATTCCTTATCATTTCTTCTTAG 1140
Db 301 AACAAATCTCTACCACTATTATTCCTTCCA-|C|C|A|A|G|T|T|C|A|T|T|A|T|C|A|T|T|G|T|C|T|T|C|T|T|A|G 359
Qy 1141 GAAACAGACATCACTCATTTGATTTGATTTAATAAAGTTTTATTTTC 1186
Db 360 GAAACAGACATGCCCAATTCATTTGATTTAATAAAGTTTTATTTTC 405

RESULT 13
US-10-027-632-120356
; Sequence 120356, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120356
; LENGTH: 970
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-120356

Query Match 17.5%; Score 219.6; DB 17; Length 970;
Best Local Similarity 81.8%; Pred. No. 1.5e-55;
Matches 332; Conservative 1; Mismatches 50; Indels 23; Gaps 6;
Qy 801 CTGGAAACCAAGAAATAGACCATCTATGAAGAATTTCTTAAGAAAGGAAACAGA 860
Db 3 CTGGAAACCAAGAAAGCCGCTCTATAAGAATTTATTAAAGAAAGGAAAGAGA 62
Qy 861 AGTTGAAAAAATCTCCCCAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCA 920
Db 63 AGTTGAAAGAACTTTCCCAAACTCAGTTGGGGCCAACTTTAATCACAGCTCCAGGACCA 122
Qy 921 GTGCAGGCTGGCTGCCCTCTTTGGCCGCGTCTGGAATATGAGCGCGCTGGCAGTCC 980
Db 123 GTGCAGGCTGGCTGCCCYC-|T|T|G|G|A|C|C|G|T|G|T|G|G|A|T|A|A|T|G|G|A|T|A|C|T|G|T|T|G|G|C|A|G|T|T|C 181
Qy 981 AGACATCAACTCAAACTGAAGCTGCAGCAATGAAGAAGCAGTCA-|-----CATAC 1031
Db 182 AGCGTCACT-CAGAACTGAAGCTGCAGCAATGAAGAAGCAGTCAATTA 240

Qy 1032 AGAAAAAGCTAATCATGC-|-----T|C|T|C|A|C|A|A|C|T|A|C|C|A|G|G|C|T|A|A|A|A|G|C|-----C 1080
Db 241 AAAAAATGTTAATCATTCCTGATGCTCTACCACTACCGTGAGGCTAAAGCAAGTC 300
Qy 1081 AAGTCAACCAACCCTATTATPACCTTCACCCAAATTCCTTATCATTTCTTCTTAG 1140
Db 301 AACAAATCTCTACCACTATTATTCCTTCCA-|C|C|A|A|G|T|T|C|A|T|T|A|T|C|A|T|T|G|T|C|T|T|C|T|T|A|G 359
Qy 1141 GAAACAGACATCACTCATTTGATTTGATTTAATAAAGTTTTATTTTC 1186
Db 360 GAAACAGACATGCCCAATTCATTTGATTTAATAAAGTTTTATTTTC 405

RESULT 14
US-10-956-157-3788/c
; Sequence 3788, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3788
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-3788

Query Match 16.4%; Score 205; DB 21; Length 567;
Best Local Similarity 79.9%; Pred. No. 3e-51;
Matches 298; Conservative 0; Mismatches 55; Indels 20; Gaps 4;
Qy 411 CCCAGGATTATGCGCGATTCAAGAAATCCATGGTGAAGGTTTGATTCCTATGAAGAA 470
Db 372 CCCAGGATTATGACGATTCCGCAATCCCTGGTGAATACTTTGGTTCTCTATGAAGTTA 313
Qy 471 AGGAGGATAAAG--TGATCAAGAGGATGCGAGCTCAGATCCGTGAGGTGGAGCAGACCG 528
Db 312 TAGAGATTGAGGTGGACCAAGGAGATGCGAGCTCAGATCTGTGAGGTGGTACAGAGCTG 253
Qy 529 ACAGAGGTGTTTGGTCTGTCTTAGAG-|-----C|C|T|C|A|G|G|C|A|G|T|C|C|C|A|G|A|C|C|C|----- 576
Db 252 GCAGGAGGTGTTTCTGCTCTCTAGAGGTTGGTTTCCCTCGAGGATCCAGACCCCTC 193
Qy 577 -----AGAAGAGGCGCTCTTCAGCACCTAGAAAGCTGGAAGGATGAACAGCAAGTAGCT 631
Db 192 AGCCAGTGTCCGGGCACTTTAGTCCCTGGAAGCTGGAAGATGAACAGCAAGTAGCT 133
Qy 632 TCCAGCTTACAGCAGCCCTCAAAATTTGGACCTGCCACAGCTCCAGAGCTTGTACTGGATG 691
Db 132 TCCAGCCACAGCAGGCGCTCCAACTTGGACCTGCCA-CAGCTCCAGAGCTTGTACTGGATG 74
Qy 692 GAGCAGACCATCTCTGACATTTCAATTGGCCATCAGGATATACAGGAGTTGGTAACATC 751
Db 73 GAGCAGGACTATCTCCGATATTCATTGGTTATCAGGATATACAGGAGTTGGTAACATC 14
Qy 752 CACTCAGTGCCA 764
Db 13 CACTCAGTGCCA 1

RESULT 15
US-10-956-157-9023
; Sequence 9023, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William

; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES

; FILE REFERENCE: 031896-043000 (AM 101081)

; CURRENT APPLICATION NUMBER: US/10/956,157

; CURRENT FILING DATE: 2004-10-04

; NUMBER OF SEQ ID NOS: 319805

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 9023

; LENGTH: 567

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-956-157-9023

Query Match 16.4%; Score 205; DB 21; Length 567;

Best Local Similarity 79.9%; Pred. No. 3e-51;

Matches 298; Conservative 0; Mismatches 55; Indels 20; Gaps 4;

```
Qy 411 CCCAGGATTATGCGGATTCAAGAAATCCATGGTGAAGGTTTGGATTTCCTATGAAGAA 470
    |||||
Db 196 CCCAGGATTATGCGGATTCCGCAATCCCTGGTGAATACTTGGTTCTCTATGAAGTTA 255

Qy 471 AGGAGGATAAAG--TGATCAAGGAGATGCGCAGCTCAGATCCGTGAGGTGGAGCAGAGCCG 528
    |||||
Db 256 TAGAGGATTGAGGTGGACCAAGGAGATGCGCAGCTCAGATCTGTGAGGTGGTACAGAGCTG 315

Qy 529 ACAGGAGTGGTTCGGTCTGTCTTAGAG-----CCTCAGGAGTGCCAGACCC---- 576
    |||||
Db 316 GCAGGAGTGGTTCAGTCTGTCTTAGAGTTGGTTTCCTCGAGGATCCAGACCCCTC 375

Qy 577 -----AGAGAGGGCTCTTACGACCTAGAGCTGGAAGGATCAACAGCCAGTAGCT 631
    |||||
Db 376 AGCCAGTGTCCGGGCACCTTTAGTCCCTGGAAGCTGGAAGGATGAACAGCCAGGTAGCT 435

Qy 632 TCCAGCTTACAGCAGCCCTCAAATTTGGACCTGCCACAGCTCCAGAGCTTTGACTGGATG 691
    |||||
Db 436 TCCAGCCACAGCAGGCTCCAACTTGGACCTGCCA-CAGCTCCAGAGCTTGACTGGATG 494

Qy 692 GAGACAGGACCATCTCTGACATTCATTGGCCCATCAGGATATACCAGGAGTTGGTAAACATC 751
    |||||
Db 495 GAGACAGGACTATCTCCGATATTCATTGGTTATCAGGATATACCAGGAGTTGGTAAACATC 554

Qy 752 CACTCAGGTGCCA 764
    |||||
Db 555 CACTCAGGTGCCA 567
```

Search completed: June 11, 2005, 06:13:08

Job time : 478.179 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 09:44:05 ; Search time 479.315 Seconds
(without alignments)
11471.453 Million cell updates/sec

Title: US-09-155-676B-4
Perfect score: 1253
Sequence: 1 cattggagtcacgcggtgacg.....ggggggcggtaccactttt 1253

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 12985247 seqs, 2194111492 residues

Total number of hits satisfying chosen parameters: 25970494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA New: *

1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq: *
2: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq: *
3: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq: *
4: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq: *
5: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq: *
6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq: *
7: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq: *
8: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq: *
9: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq: *
10: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq: *
11: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq: *
12: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq: *
13: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq: *
14: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq: *
15: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1252	99.9	1253	6	US-09-155-676B-4
2	1252	99.9	1253	13	US-11-060-756-1553
3	1252	99.9	1253	13	US-11-060-756-1554
4	1252	99.9	1253	13	US-11-060-756-1554
5	1252	99.9	1253	13	US-11-060-756-5825
6	514.6	41.1	681	8	US-10-450-763-19571
7	479.6	38.3	2557	8	US-10-450-763-19571
8	370.4	29.6	742	8	US-10-450-763-19570
9	323.6	25.8	743	15	US-60-680-544-781
10	323.6	25.8	743	15	US-60-680-544-7415
11	323.6	25.8	743	15	US-60-680-473-781
12	323.6	25.8	743	15	US-60-680-473-7415
13	278.4	22.2	816	11	US-10-940-774A-3437
14	263.4	21.0	3100	8	US-10-450-763-19572
15	215.2	17.2	4017	8	US-10-450-763-19572
16	175.8	14.0	6623	11	US-10-940-774A-16221
17	175.8	14.0	7031	11	US-10-940-774A-15179
18	175.8	14.0	246960	12	US-11-121-086-8

19 132.8 10.6 1600 8 US-10-450-763-13672 Sequence 13672, A
20 130.8 10.4 467 8 US-10-450-763-19569 Sequence 19569, A
21 104.8 8.4 258 8 US-10-006-290-8756 Sequence 8756, Ap
22 104.8 8.4 258 10 US-10-990-275-8756 Sequence 8756, Ap
23 104.8 8.4 258 10 US-10-990-298-8756 Sequence 8756, Ap
24 76 6.1 166 15 US-60-680-544-30183 Sequence 30183, A
25 76 6.1 166 15 US-60-680-473-30183 Sequence 30183, A
26 61.6 4.9 1906 6 US-09-155-676B-1 Sequence 1, Appli
27 53.2 4.2 2631 6 US-09-155-676B-3 Sequence 3, Appli
28 43.8 3.5 1981 6 US-09-783-931A-26 Sequence 26, Appli
29 43.4 3.5 224 13 US-11-021-493A-335 Sequence 335, App
30 40.8 3.3 1169 9 US-10-703-032-17036 Sequence 17036, A
31 40.8 3.3 5521 8 US-10-450-763-7353 Sequence 7353, Ap
32 40.8 3.3 62267 11 US-10-990-328A-94216 Sequence 94216, A
33 40.8 3.3 71278 11 US-10-940-774A-11851 Sequence 11851, A
34 40.8 3.2 71278 11 US-10-940-774A-11851 Sequence 17563, A
35 40.4 3.2 3738 2 PCT-US05-02325-59 Sequence 59, Appl
36 40.4 3.2 3803 2 PCT-US05-02325-55 Sequence 55, Appl
37 40.4 3.2 3803 15 US-60-664-611-2372 Sequence 2372, Ap
38 40.4 3.2 3805 15 US-60-659-397-549 Sequence 549, App
39 40.4 3.2 3834 15 US-60-659-397-548 Sequence 548, App
40 40.4 3.2 3840 15 US-60-685-372-563 Sequence 563, App
41 40.4 3.2 3943 11 US-10-940-774A-2495 Sequence 2495, Ap
42 40.4 3.2 3943 11 US-10-940-774A-2496 Sequence 2496, Ap
43 40.4 3.2 3943 11 US-10-940-774A-2497 Sequence 2497, Ap
44 40.4 3.2 3945 15 US-60-659-397-550 Sequence 550, App
45 40.4 3.2 3952 15 US-60-664-611-2374 Sequence 2374, Ap

ALIGNMENTS

RESULT 1

US-09-155-676B-4
; Sequence 4, Application US/09155676B
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR
; FILE OF INVENTION: PREPARATION AND USE
; FILE REFERENCE: WALLACH=21
; CURRENT APPLICATION NUMBER: US/09/155,676B
; CURRENT FILING DATE: 1999-01-04
; PRIOR APPLICATION NUMBER: PCT/IL97/00117
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: IL 117800
; PRIOR FILING DATE: 1996-04-02
; PRIOR APPLICATION NUMBER: IL 119133
; PRIOR FILING DATE: 1996-08-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 4
; LENGTH: 1253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: n is a, c, g, or t
US-09-155-676B-4

Query Match 99.9%; Score 1252; DB 6; Length 1253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CATTGGAGTCACGGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCANGAATTC 60
Db 1 CATTGGAGTCACGGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCANGAATTC 60
Qy 61 GATTGCGACCCACGAAGGCCCTTTCTTCTGTGTGTCGGGCACGTTTATAGCCGCAAGCAC 120

Db 61 GATTGAGCCAGGAAGGCCCTTCTCTGTGGTCCGCGCAGCTTTACAGCCGCAAGCAC 120
Qy 121 CCAGCGGAGCTGAAGGAGCTTTTGAAGGCTCTCTGCCAGAGTGGAGCGGCCGCAA 180
Db 121 CCAGCGGAGCTGAAGGAGCTTTTGAAGGCTCTCTGCCAGAGTGGAGCGGCCGCAA 180
Qy 181 GGCATCCGCGCGCTCAGGTGGAGCTTATGTGCCGAAACACAGAGCGATGCTGTGGTG 240
Db 181 GGCATCCGCGCGCTCAGGTGGAGCTTATGTGCCGAAACACAGAGCGATGCTGTGGTG 240
Qy 241 CTGTGCTGCGCTGTGAGTGGGGAACACCTGAGCCATGAAACCTGACGGTCTGTA 300
Db 241 CTGTGCTGCGCTGTGAGTGGGGAACACCTGAGCCATGAAACCTGACGGTCTGTA 300
Qy 301 CGGGGGCTGCTGGAGCTCTGGCAGCCAGAGCACAGAAAGCAACCAAAATTCG 360
Db 301 CGGGGGCTGCTGGAGCTCTGGCAGCCAGAGCACAGAAAGCAACCAAAATTCG 360
Qy 361 GTGGGAGAAACAAAGCTGAGTCCAGATGAAAGAGATTTCTGTCTACTCCCGAGATTA 420
Db 361 GTGGGAGAAACAAAGCTGAGTCCAGATGAAAGAGATTTCTGTCTACTCCCGAGATTA 420
Qy 421 TGGCGGATTCAGAAATCATCGTGAAGTTTGGATTCCTATGAAGAAAGAGATTA 480
Db 421 TGGCGGATTCAGAAATCATCGTGAAGTTTGGATTCCTATGAAGAAAGAGATTA 480
Qy 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGT 540
Db 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGT 540
Qy 541 TCGGCTGTCTTAGAGCTTCAGGAGTGCAGACCCAGAGAGGCTCTTCAGACCTAG 600
Db 541 TCGGCTGTCTTAGAGCTTCAGGAGTGCAGACCCAGAGAGGCTCTTCAGACCTAG 600
Qy 601 AAGCTGGAAGGATGAAACAGCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGA 660
Db 601 AAGCTGGAAGGATGAAACAGCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGA 660
Qy 661 CTGCGCAGCTCCAGAGCTTGAAGTGGATGAGAGCAGGACCATCTCTGACATTCATGG 720
Db 661 CTGCGCAGCTCCAGAGCTTGAAGTGGATGAGAGCAGGACCATCTCTGACATTCATGG 720
Qy 721 CCATCAGGATATACAGGAGTTGGTAAACATCCACTCAGGTGCCACACCTCCCTGGATGAT 780
Db 721 CCATCAGGATATACAGGAGTTGGTAAACATCCACTCAGGTGCCACACCTCCCTGGATGAT 780
Qy 781 CCAAGATGAAGATACATTTGCTGGGAAACCAAGAAATAGGACCATCTATGAAGAAATTTCT 840
Db 781 CCAAGATGAAGATACATTTGCTGGGAAACCAAGAAATAGGACCATCTATGAAGAAATTTCT 840
Qy 841 TAAAGAAAGGAAACAGAGCTTGAAGAACTCCCGCCAGAGCTGGGCGCAACTT 900
Db 841 TAAAGAAAGGAAACAGAGCTTGAAGAACTCCCGCCAGAGCTGGGCGCAACTT 900
Qy 901 TGATCACAGCTCCAGGACCAAGTGGAGCTGGCTGCTCTTTTGGCGCGCTCTGAATA 960
Db 901 TGATCACAGCTCCAGGACCAAGTGGAGCTGGCTGCTCTTTTGGCGCGCTCTGAATA 960
Qy 961 ATGGACGCGCTGGGAGTCCAGACATCAACTCCAAAACCTGAAGCTGCAGCAATGAAGAAG 1020
Db 961 ATGGACGCGCTGGGAGTCCAGACATCAACTCCAAAACCTGAAGCTGCAGCAATGAAGAAG 1020
Qy 1021 CAGTCACATACAGAAAGCTTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCC 1080
Db 1021 CAGTCACATACAGAAAGCTTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCC 1080
Qy 1081 AAAGTCAACCAACCCCTATTATACCTTCCACCAAAATCTTTATCATTTCTTTCTTAG 1140
Db 1081 AAAGTCAACCAACCCCTATTATACCTTCCACCAAAATCTTTATCATTTCTTTCTTAG 1140
Qy 1141 GAAACAGACATACATTCATTTGATTTAATAAGTTTATTTTTCGGCCTTCGTGGCCT 1200
Db 1141 GAAACAGACATACATTCATTTGATTTAATAAGTTTATTTTTCGGCCTTCGTGGCCT 1200

Db 1141 GAAACAGACATACATTCATTTGATTTAATAAGTTTATTTTTCGGCCTTCGTGGCCT 1200
Qy 1201 CGAATCAAGCTTATCGATACCGTGCAGCTCGAGGGGGCGGTACCCACTTTT 1253
Db 1201 CGAATCAAGCTTATCGATACCGTGCAGCTCGAGGGGGCGGTACCCACTTTT 1253
RESULT 2
US-11-060-756-1553
; Sequence 1553, Application US/11060756
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1553
; LENGTH: 1253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: n is a, c, g, or t
US-11-060-756-1553
Query Match 99.9%; Score 1252; DB 13; Length 1253;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CATTGGAGTACAGCGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60
Db 1 CATTGGAGTACAGCGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60
Qy 61 GATTGAGCCACGAAGGCCCTTCTCTGTGTGGTGGCGGACGTTTACAGCCGCAAGCAC 120
Db 61 GATTGAGCCACGAAGGCCCTTCTCTGTGTGGTGGCGGACGTTTACAGCCGCAAGCAC 120
Qy 121 CCAGCGGAGCTGAAGGAGGCTTTTGAGAGCTCTCTGCCAGAGTGGAGCGGCCGCAA 180
Db 121 CCAGCGGAGCTGAAGGAGGCTTTTGAGAGCTCTCTGCCAGAGTGGAGCGGCCGCAA 180
Qy 181 GGCATCCGCGCGCTCAGGTGGAGCGCTATGTGCCGAAACACAGAGCGATGCTGTGGTG 240
Db 181 GGCATCCGCGCGCTCAGGTGGAGCGCTATGTGCCGAAACACAGAGCGATGCTGTGGTG 240
Qy 241 CTGTGCTGCGCTGTGAGGTGGGAAACACCTGAGCCATGGAACCTGACGGTCTGTA 300
Db 241 CTGTGCTGCGCTGTGAGGTGGGAAACACCTGAGCCATGGAACCTGACGGTCTGTA 300
Qy 301 CGGGGGCTGCTGGAGCATCTGGCCAGCCAGAGCACAAAGAAAGCAACCAAAATTCG 360
Db 301 CGGGGGCTGCTGGAGCATCTGGCCAGCCAGAGCACAAAGAAAGCAACCAAAATTCG 360
Qy 361 GTGGGAGAAACAAAGCTGAGTCCAGATGAAAGAGATTTCTGTCTACTCCCGAGATTA 420
Db 361 GTGGGAGAAACAAAGCTGAGTCCAGATGAAAGAGATTTCTGTCTACTCCCGAGATTA 420
Qy 421 TGGCGGATTCAGAAATCATCGTGAAGTTTGGATTCCTATGAAGAAAGAGGATTA 480
Db 421 TGGCGGATTCAGAAATCATCGTGAAGTTTGGATTCCTATGAAGAAAGAGGATTA 480
Qy 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGT 540
Db 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGT 540
Qy 541 TCGGCTGTCTTAGAGCTTCAGGAGTGCAGACCCAGAGAGGCTCTTCAGACCTAG 600
Db 541 TCGGCTGTCTTAGAGCTTCAGGAGTGCAGACCCAGAGAGGCTCTTCAGACCTAG 600


```
QY 601 AAGCTGGAAGAGGATGAACAGCCAAAGTAGCTTTCAGCTTACAGCAGCCCTCAAATTTGGA 660
Db 601 AAGCTGGAAGAGGATGAACAGCCAAAGTAGCTTTCAGCTTACAGCAGCCCTCAAATTTGGA 660
QY 661 CTTGCCACAGCTCCAGAGCTTGACTGATGGAGACAGGACCATCTGACATTCATTGG 720
Db 661 CTTGCCACAGCTCCAGAGCTTGACTGATGGAGACAGGACCATCTGACATTCATTGG 720
QY 721 CCATCAGGATATACCAAGGATTTGTAACATCCACTCAGTGCCACACCTCCCTGGATGAT 780
Db 721 CCATCAGGATATACCAAGGATTTGTAACATCCACTCAGTGCCACACCTCCCTGGATGAT 780
QY 781 CCAAGATGAAGATATACATTTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCT 840
Db 781 CCAAGATGAAGATATACATTTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCT 840
QY 841 TAAAGAAAGGAAAGAAAGAGAGTTGAAAGAACTCCCCCCAGACCGAGTTGGGGCCAACTT 900
Db 841 TAAAGAAAGGAAAGAAAGAGAGTTGAAAGAACTCCCCCCAGACCGAGTTGGGGCCAACTT 900
QY 901 TGATCACAGCTCCAGGACCAAGTGCAGGCTGGCTGCCCTCTTTTGGGGCGCGCTCTGGAATA 960
Db 901 TGATCACAGCTCCAGGACCAAGTGCAGGCTGGCTGCCCTCTTTTGGGGCGCGCTCTGGAATA 960
QY 961 ATGAGCGCGCTGGCAGTCCAGACATCAATCCAAACTGAAAGCTGACGCAATGAAGAAG 1020
Db 961 ATGAGCGCGCTGGCAGTCCAGACATCAATCCAAACTGAAAGCTGACGCAATGAAGAAG 1020
QY 1021 CAGTCACATACAGAAAGAAAGCTATCATGCTCTTACCAACTACCATGAGGCTAAAGCC 1080
Db 1021 CAGTCACATACAGAAAGAAAGCTATCATGCTCTTACCAACTACCATGAGGCTAAAGCC 1080
QY 1081 AAAGTCAACCAACCCCTATTATACCTTCCACCCAAATTTCTTATCATTTCTTTCTAG 1140
Db 1081 AAAGTCAACCAACCCCTATTATACCTTCCACCCAAATTTCTTATCATTTCTTTCTAG 1140
QY 1141 GAAACAGACATACCTCATTTGATTTAATAAGTTTTATTTTTCGGCCCTTCGTCGCT 1200
Db 1141 GAAACAGACATACCTCATTTGATTTAATAAGTTTTATTTTTCGGCCCTTCGTCGCT 1200
QY 1201 CGAATCAGCTTATCGATACCGTCAGCTCGACCTCGAGGGGGCGGTACCCACTTTT 1253
Db 1201 CGAATCAGCTTATCGATACCGTCAGCTCGACCTCGAGGGGGCGGTACCCACTTTT 1253
```

RESULT 3

```
US-11-060-756-1554
; Sequence 1554, Application US/11060756
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1554
; LENGTH: 1253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: n is a, c, g, or t
US-11-060-756-1554
```

Query Match 99.9%; Score 1252; DB 13; Length 1253;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CATTGGAGTCACGGGTGGCGCGCTCTAGATAGTGGATCCCGGGCTGCANGGAATTC 60
Db 1 CATTGGAGTCACGGGTGGCGCGCTCTAGATAGTGGATCCCGGGCTGCANGGAATTC 60
QY 61 GATTTCGAGCCACCAAGGCCCTTTCTTCTGTGTGTCGGGACAGCTTTTACAGCCGCAAGCAC 120
Db 61 GATTTCGAGCCACCAAGGCCCTTTCTTCTGTGTGTCGGGACAGCTTTTACAGCCGCAAGCAC 120
QY 121 CCAGCGCAGCTGAAGAGGCGCTTTTGAGAGGCTCTCTGCCCCAGGTGGAGGGCGGCCGCA 180
Db 121 CCAGCGCAGCTGAAGAGGCGCTTTTGAGAGGCTCTCTGCCCCAGGTGGAGGGCGGCCGCA 180
QY 181 GGCCATCCGCGCCCTCAGGTGGAGCGCTATGTGCCCGAAACAGGCGATCTCTGCTGTG 240
Db 181 GGCCATCCGCGCCCTCAGGTGGAGCGCTATGTGCCCGAAACAGGCGATCTCTGCTGTG 240
QY 241 CTTGTGCTGCGGCTGTGAGGTGCGGAAACACTCAGGCCATCGAAACCTTGACGGTCTCTGA 300
Db 241 CTTGTGCTGCGGCTGTGAGGTGCGGAAACACTCAGGCCATCGAAACCTTGACGGTCTCTGA 300
QY 301 CGGGGGCTGCTGGAGCATCTGGCCAGCCAGAGCAAAAGAAAGCAACCAAAATTTCTG 360
Db 301 CGGGGGCTGCTGGAGCATCTGGCCAGCCAGAGCAAAAGAAAGCAACCAAAATTTCTG 360
QY 361 GTGGAGAAACAAAGCTCAGGTCCAGATGAAAGAAAGTTTCTGTCTCACTCCCAAGATTA 420
Db 361 GTGGAGAAACAAAGCTCAGGTCCAGATGAAAGAAAGTTTCTGTCTCACTCCCAAGATTA 420
QY 421 TGCCCGATTCAAGAAATCCATGTGTGAAAGTTTCTGATTTCTATGAAGAAAGAGGATTA 480
Db 421 TGCCCGATTCAAGAAATCCATGTGTGAAAGTTTCTGATTTCTATGAAGAAAGAGGATTA 480
QY 481 AGTGATCAAGAGATGCGAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGGAGTGGT 540
Db 481 AGTGATCAAGAGATGCGAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGGAGTGGT 540
QY 541 TCGTCTGTCTTAGAGCTCAGGAGTCCAGACCCAGAGAGGGCTCTTTCAGCACTTAG 600
Db 541 TCGTCTGTCTTAGAGCTCAGGAGTCCAGACCCAGAGAGGGCTCTTTCAGCACTTAG 600
QY 601 AAGCTGGAAGGAGTGAACAGCCAAAGTAGCTTCCAGCTTACAGCAGCCCTCAAATTTGGA 660
Db 601 AAGCTGGAAGGAGTGAACAGCCAAAGTAGCTTCCAGCTTACAGCAGCCCTCAAATTTGGA 660
QY 661 CCTGCCACAGCTCCAGAGCTTGA CTGATGGAGACAGGACCATCTCTGACATTCATTGG 720
Db 661 CCTGCCACAGCTCCAGAGCTTGA CTGATGGAGACAGGACCATCTCTGACATTCATTGG 720
QY 721 CCATCAGGATATACCAAGGATTTGTAACATCCACTCAGGTGCCACACCTCCCTGGATGAT 780
Db 721 CCATCAGGATATACCAAGGATTTGTAACATCCACTCAGGTGCCACACCTCCCTGGATGAT 780
QY 781 CCAAGATGAAGATATACATTTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCT 840
Db 781 CCAAGATGAAGATATACATTTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCT 840
QY 841 TAAAGAAAGGAAAGAAAGAGTTGAAAGAACTCCCCCCAGACCGAGTTGGGGCCAACTT 900
Db 841 TAAAGAAAGGAAAGAAAGAGTTGAAAGAACTCCCCCCAGACCGAGTTGGGGCCAACTT 900
QY 901 TGATCAGCTCCAGGACCAAGTGCAGGCTGGCTGCCCTCTTTTGGGGCGCGCTCTGGAATA 960
Db 901 TGATCAGCTCCAGGACCAAGTGCAGGCTGGCTGCCCTCTTTTGGGGCGCGCTCTGGAATA 960
QY 961 ATGAGCGCGCTGGCAGTCCAGACATCAATCCAAACTGAAAGCTGACGCAATGAAGAAG 1020
Db 961 ATGAGCGCGCTGGCAGTCCAGACATCAATCCAAACTGAAAGCTGACGCAATGAAGAAG 1020
QY 1021 CAGTCACATACAGAAAGAAAGCTATCATGCTCTTACCAACTACCATGAGGCTAAAGCC 1080
Db 1021 CAGTCACATACAGAAAGAAAGCTATCATGCTCTTACCAACTACCATGAGGCTAAAGCC 1080
QY 1081 AAAGTCAACCAACCCCTATTATACCTTCCACCCAAATTTCTTATCATTTCTTTCTAG 1140
```

Db 1081 AAAGTCAACAAACCCCTATTATACCTTCCACCAAAATTTCTTTATCATTTGCTTCTTAG 1140
QY 1141 GAAACAGACATACCTATTGATTTGATTTAAATTAAGTTTATTTTCGGCCTTCGTGGCCT 1200
Db 1141 GAAACAGACATACCTATTGATTTGATTTAAATTAAGTTTATTTTCGGCCTTCGTGGCCT 1200
QY 1201 CGAATCAAGCTTATCGATACCTCGACCTCGAGGGGGGGCCGTACCCACTTTT 1253
Db 1201 CGAATCAAGCTTATCGATACCTCGACCTCGAGGGGGGGCCGTACCCACTTTT 1253

RESULT 4
US-11-060-756-5825
; Sequence 5825, Application US/11060756
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5825
; LENGTH: 1253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: n is a, c, g, or t
US-11-060-756-5825

Query Match 99.9%; Score 1252; DB 13; Length 1253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATTGGAGTCACGGTGGCGGCTCTAGATAGTGGATCCCGGCTGCANGGAATTC 60
Db 1 CATTGGAGTCACGGTGGCGGCTCTAGATAGTGGATCCCGGCTGCANGGAATTC 60
QY 61 GATTGAGCCACGAAGGCCCTTCTTCTGTGGTGGCGGACGTTTACAGCCGCAAGCAC 120
Db 61 GATTGAGCCACGAAGGCCCTTCTTCTGTGGTGGCGGACGTTTACAGCCGCAAGCAC 120
QY 121 CAGCGGACGCTGAAGGAGGCTTTTGAAGGCTCTCTGCCCCAGGTGGAGGGCCCGCAA 180
Db 121 CAGCGGACGCTGAAGGAGGCTTTTGAAGGCTCTCTGCCCCAGGTGGAGGGCCCGCAA 180
QY 181 GGCCATCCGCGCGCTCAGGTGGAGGCTATGTGCCGACGAGCGATGCTGTGGT 240
Db 181 GGCCATCCGCGCGCTCAGGTGGAGGCTATGTGCCGACGAGCGATGCTGTGGT 240
QY 241 CTGTGCTCGGCTGTGAGTGGCGGACACCTCGAGCCATGAAACCTGACGGTCTGTA 300
Db 241 CTGTGCTCGGCTGTGAGTGGCGGACACCTCGAGCCATGAAACCTGACGGTCTGTA 300
QY 301 CGGGGGGCTGTGGAGATCTGTGGCAGCCGACAGCAAGAAAGCAACCAAAATTCG 360
Db 301 CGGGGGGCTGTGGAGATCTGTGGCAGCCGACAGCAAGAAAGCAACCAAAATTCG 360
QY 361 GTGGGAGACAAAGCTGAGTCCAGTGAAGAGAGTTCTGTCACCTCCCGAGATTA 420
Db 361 GTGGGAGACAAAGCTGAGTCCAGTGAAGAGAGTTCTGTCACCTCCCGAGATTA 420
QY 421 TCGCGGATTCAAGAAATCATGTTGAAAGTTTGGATTCTTATGAAGAAAGAGATTA 480
Db 421 TCGCGGATTCAAGAAATCATGTTGAAAGTTTGGATTCTTATGAAGAAAGAGATTA 480
QY 481 AGTGATCAAGGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGCCGACGAGGTGT 540

Db 481 AGTGATCAAGGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGACCGCAGAGGTGT 540
QY 541 TCGGTCTGTCTTAGAGCCTCAGGAGTGGCAGACCCAGAGAGGGCTCTTTCAGCACCTAG 600
Db 541 TCGGTCTGTCTTAGAGCCTCAGGAGTGGCAGACCCAGAGAGGGCTCTTTCAGCACCTAG 600
QY 601 AAGCTGGAAAGGGATGAACAGCAAGTAGTCTTCCAGCTTACAGCAGCCCTCAAAATTGGA 660
Db 601 AAGCTGGAAAGGGATGAACAGCAAGTAGTCTTCCAGCTTACAGCAGCCCTCAAAATTGGA 660
QY 661 CCTGCCACAGCTCCAGAGCTTGAAGTGGAGAGAGAGCCATCTCTGACATTCATTGG 720
Db 661 CCTGCCACAGCTCCAGAGCTTGAAGTGGAGAGAGAGCCATCTCTGACATTCATTGG 720
QY 721 CCATCAGGATATACCAGGAGTTGGTAAATCAATCCACTCAGGTGCGCACACCTCCCTGGATGAT 780
Db 721 CCATCAGGATATACCAGGAGTTGGTAAATCAATCCACTCAGGTGCGCACACCTCCCTGGATGAT 780
QY 781 CCAAGATGAAGAATAATCTTGGTGGGAAACCAAGAAATAGACCATCTATGAAGAAATTTCT 840
Db 781 CCAAGATGAAGAAATATCTTGGTGGGAAACCAAGAAATAGACCATCTATGAAGAAATTTCT 840
QY 841 TAAAGAAAGGAAAGAAACAGAAAGTTGAAAGAACTCCCCCAGACCCAGTTGGGGCCAACTT 900
Db 841 TAAAGAAAGGAAAGAAACAGAAAGTTGAAAGAACTCCCCCAGACCCAGTTGGGGCCAACTT 900
QY 901 TGATCACAGCTCCAGGACGAGTGGAGTGGCTGCTCTTTTGGGGCCGCTCTGGAATA 960
Db 901 TGATCACAGCTCCAGGACGAGTGGAGTGGCTGCTCTTTTGGGGCCGCTCTGGAATA 960
QY 961 ATGACAGCGCTGGCAGTCCAGACATCAATCCAAATCTGAAAGCTGACGCAATGAAGAAG 1020
Db 961 ATGACAGCGCTGGCAGTCCAGACATCAATCCAAATCTGAAAGCTGACGCAATGAAGAAG 1020
QY 1021 CAGTCACATACAGAAAAAGCTTAATCATGCTCTCTCAACATACATGAGGCTAAAGCC 1080
Db 1021 CAGTCACATACAGAAAAAGCTTAATCATGCTCTCTCAACATACATGAGGCTAAAGCC 1080
QY 1081 AAAGTCAACCAACCCCTATTATACCTTCCACCAAAATTTCTTATCATTTGCTTCTTAG 1140
Db 1081 AAAGTCAACCAACCCCTATTATACCTTCCACCAAAATTTCTTATCATTTGCTTCTTAG 1140
QY 1141 GAAACAGACATACCTATTGATTTGATTTAAATTAAGTTTATTTTCGGCCTTCGTGGCCT 1200
Db 1141 GAAACAGACATACCTATTGATTTGATTTAAATTAAGTTTATTTTCGGCCTTCGTGGCCT 1200
QY 1201 CGAATCAAGCTTATCGATACCTCGACCTCGAGGGGGGGCCGTACCCACTTTT 1253
Db 1201 CGAATCAAGCTTATCGATACCTCGACCTCGAGGGGGGGCCGTACCCACTTTT 1253

RESULT 5
US-11-060-756-5826
; Sequence 5826, Application US/11060756
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5826
; LENGTH: 1253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: n is a, c, g, or t
US-11-060-756-5826

Query Match 99.9%; Score 1252; DB 13; Length 1253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATTGGAGTCACGGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGANGGAATTC 60
Db 1 CATTGGAGTCACGGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGANGGAATTC 60

Qy 61 GATTTCGAGCCCAAGAGCCCTCTCTCTGGTGGTGGCGACAGTTTACAGCCGCAAGCAC 120
Db 61 GATTTCGAGCCCAAGAGCCCTCTCTCTGGTGGTGGCGACAGTTTACAGCCGCAAGCAC 120

Qy 121 CCAGCGCAGCTGAAGAGGCTTTTGAAGAGCTCTCTGCCCCAGGTGGAGGCGGCCCGCAA 180
Db 121 CCAGCGCAGCTGAAGAGGCTTTTGAAGAGCTCTCTGCCCCAGGTGGAGGCGGCCCGCAA 180

Qy 181 GGCCATCCGCGCGCTCAGGTGGAGGCTATGTCGCCGACACGAGCGATCTCTGGTG 240
Db 181 GGCCATCCGCGCGCTCAGGTGGAGGCTATGTCGCCGACACGAGCGATCTCTGGTG 240

Qy 241 CCTGTGCTGCGGCTGTGAGGTGGCGGACACCTGAGCCATCGAAACCTGACGGTGTGTA 300
Db 241 CCTGTGCTGCGGCTGTGAGGTGGCGGACACCTGAGCCATCGAAACCTGACGGTGTGTA 300

Qy 301 CGGGGGGCTGTGAGCATCTGGCCAGCCAGAGCACAAGAAAGCAACCAAAATCTG 360
Db 301 CGGGGGGCTGTGAGCATCTGGCCAGCCAGAGCACAAGAAAGCAACCAAAATCTG 360

Qy 361 GTGGAGNACAAAGCTGAGGTCCAGATCAAGAGAGTTTCTGTCTCACTCCAGGATTA 420
Db 361 GTGGAGNACAAAGCTGAGGTCCAGATCAAGAGAGTTTCTGTCTCACTCCAGGATTA 420

Qy 421 TGCCGATTCAAGAAATCCATGTTGAAAGTTTCTGATTCCTATCAAGAAAGGAGGATAA 480
Db 421 TGCCGATTCAAGAAATCCATGTTGAAAGTTTCTGATTCCTATCAAGAAAGGAGGATAA 480

Qy 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGGAGTGGT 540
Db 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGGAGTGGT 540

Qy 541 TCGTCTGTCTTAGAGCTCAGGAGTGCAGACCCAGAGAGGCGTCTTCAGCACCTAG 600
Db 541 TCGTCTGTCTTAGAGCTCAGGAGTGCAGACCCAGAGAGGCGTCTTCAGCACCTAG 600

Qy 601 AAGCTGGAAGGGATGAACAGCCAGTAGCTTCAGCTTACAGCAGCCCTCAAAATTTGGA 660
Db 601 AAGCTGGAAGGGATGAACAGCCAGTAGCTTCAGCTTACAGCAGCCCTCAAAATTTGGA 660

Qy 661 CCTGCCACAGCTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGG 720
Db 661 CCTGCCACAGCTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGG 720

Qy 721 CCATCAGGATATACAGAGTTGGTAAATCACTCAGGTGGCCACACCTCTCTGGATGAT 780
Db 721 CCATCAGGATATACAGAGTTGGTAAATCACTCAGGTGGCCACACCTCTCTGGATGAT 780

Qy 781 CCAAGATGAAGATACATCTGCGGAACCAAGAAATAGGACCATCTATGAAGATTTCT 840
Db 781 CCAAGATGAAGATACATCTGCGGAACCAAGAAATAGGACCATCTATGAAGATTTCT 840

Qy 841 TAAAGAAAGGAAAAACAGAAATTTGAAAAAACTCCCGCCAGACCGAGTTGGGGCAACTT 900
Db 841 TAAAGAAAGGAAAAACAGAAATTTGAAAAAACTCCCGCCAGACCGAGTTGGGGCAACTT 900

Qy 901 TGATCAGCTCCAGGACCAAGTGCAGGTGGCTGCCCTCTTTTGGGCGCGTCTGGAATA 960
Db 901 TGATCAGCTCCAGGACCAAGTGCAGGTGGCTGCCCTCTTTTGGGCGCGTCTGGAATA 960

Qy 961 ATGACGCGCTGCGAGTCCAGATCAATCCCAAAACTGAGCTGAGCAGCATGAAGAAG 1020
Db 961 ATGACGCGCTGCGAGTCCAGATCAATCCCAAAACTGAGCTGAGCAGCATGAAGAAG 1020

RESULT 6

US-10-450-763-19571
; Sequence 19571, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 19571
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (187)..(675)
; OTHER INFORMATION: 78% homologous to Homo sapiens TRAF2 binding protein encoded
; OTHER INFORMATION: by clone 15, accession number W42401, Smith-Waterman Score=665.
US-10-450-763-19571

Query Match 41.1%; Score 514.6; DB 8; Length 681;
Best Local Similarity 97.4%; Pred. No. 8.6e-124;
Matches 568; Conservative 0; Mismatches 4; Indels 11; Gaps 4;

Qy 81 CCTTCTCTCTGTGGTGGC-GGCACGTTTACAGCCGCAAGCAGCCAGCGCAGCTCAAGGAG 139
Db 99 CCTTCTCTCTGTGGTGGCAGCGTTCAGCGCCGCAAGCA-CCAGCGCAGCTCAAGGAG 157

Qy 140 GCTTTTGAAGGCTCTCTGCCCC-----AGTGGAGGCGGCGCCGCAAGCCATCCGCG 191
Db 158 GCTTTGAGAGGCTCTCTGCCCCAGGTGGAGGCGGCGCCGCAAGCCATCCGCG 217

Qy 192 CCCTCAGGTGGAGCGCTATGTCGCCGAAACAGCGCATGCTGCTGCTGCTGCTGCTGCG 251
Db 218 CCCTCAGGTGGAGCGCTATGTCGCCGAAACAGCGCATGCTGCTGCTGCTGCTGCTGCG 277

Qy 252 GCTGTGAGGTGGCGGAACACCTGAGCCATGGAACCTGACGCTGCTGTACGGGGGGTGC 311
Db 278 GCTGTGAGGTGGCGGAACACCTGAGCCATGGAACCTGACGCTGCTGTACGGGGGGTGC 337

Qy 312 TGGAGCATCTGGCCAGCCAGAGCACAAGAAAGCAACCAAAATCTTGTGGGAGAAC 371
Db 338 TGGAGCATCTGGCCAGCCAGAGCACAAGAAAGCAACCAAAATCTTGTGGGAGAAC 397

Qy 372 AAGCTGAGGTCCAGATGAAGAGAGTTTCTGTGCTCACTCCAGGATTTATGCGGATTC 431
Db 398 AAGCTGAGGTCCAGATGAAGAGAGTTTCTGTGCTCACTCCAGGATTTATGCGGATTC 457

QY 432 AGAATCCATGTTGAAGGTTGGATTCTTATGAAGAAAGAGGATAAAGTGATCAAGG 491
Db 458 AGAAATCCATGTTGAAGGTTGGATTCTTATGAAGAAAGAGGATAAAGTGATCAAGG 517
QY 492 AGATGCGAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACA-GGAGGTGTTTGGTCTGTC 550
Db 518 AGATGCGAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACA-GGAGGTGTTTGGTCTGTC 577
QY 551 TTAGAGCCTCAGGAGTGCAGACCCAGAGAGGGCTCTTACGACACTAGAAAGCTGGAAA 610
Db 578 TTAGAGCCTCAGGAGTGCAGACCCAGAGAGGGCTCTTACGACACTAGAAAGCTGGAAA 637
QY 611 GGGATGAACAGCAAGTACTCTCCAGCTTACAGCAGCCCTCAA 653
Db 638 GGGATGAACAGCAAGTACTCTTACGCTTACAGCAGCCCTTAA 680

RESULT 7

US-10-450-763-23046
; Sequence 23046, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 23046
; LENGTH: 2557
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1885)..(2196)
; OTHER INFORMATION: 40% homologous to Mus musculus synbindin, accession number AF233340, Smith-Waterman Score=116.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2557)
; OTHER INFORMATION: n = a,t,c or g
US-10-450-763-23046

Query Match 38.3%; Score 479.6; DB 8; Length 2557;
Best Local Similarity 84.9%; Pred. No. 1.8e-114;
Matches 702; Conservative 0; Mismatches 55; Indels 70; Gaps 12;

QY 426 GATTCAGAAATCCATGTTGAAGGTTGGATTCTTATGAAGAAAGAGGATAAAGTGA 485
Db 36 GATTCAGAAATCCATGTTGAAGGTTGGATTCTTATGAAGAAAGAGGATAAAGTGA 95
QY 486 TCAAGGAGATGCGAGTCCGTGAGGTGGAGCAGAGCCGACAGGAGGTGTTCCGT 545
Db 96 TCAAGGAGATGCGAGTCCGTGAGGTGGAGCAGAGCCGACAGGAGGTGTTCCGT 155
QY 546 CTGTCTTTAGAG-----CCTCAGGCGAGTGCAGACCCA 577
Db 156 CTGTCTTTAGAGTTGTTTCCCTCGGAGGATCCAGACCACTCAGGCGAGTGCAGACCCA 215
QY 578 GAAGAGGGCTCTTCAGACCTA-GAAGCTGGAAGGATGAACGCCAGTAGCTTT-CCA 635
Db 216 GAAGAGGGCTCTTCAGACCTAAGAGGATGAACGCCAGTAGCTTTCCCA 275
QY 636 GCTTACAGCAGCCCTCAAAATTTGGACCTGCCACACGAGCTCCAGAGCTTGATGGAGGA 695
Db 276 GCTTACAGCAGCCCTCAAAATTTGGACCTGCCACACGAGCTCCAGAGCTTGATGGAGGA 335

QY 696 CAGGACCATCTCTGTGACATTTCAATTGGCCCATCAGGATATA-CCAGGAGTTGGTAAACATCCAC 754
Db 336 CAGGACCATCTCTGTGACATTTCAATTGGCCCATCAGGATATA-CCAGGAGTTGGTAAACATCCAC 395
QY 755 TCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATACAT-----TGCTGGGAACCA 810
Db 396 TCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATACATTTGGCTGGGGGACCCCA 455
QY 811 AGAAATAGGACC--ATCCTATGAAGAATTTCTTAAAGAAA-----848
Db 456 AGAAATNGGGCCCATCTCCCTATGAAGAATTTCTTTAAAGAAATTTCCCTCATCATTTGTAC 515
QY 849 --AGGAAAAACAGAAAGTTGAAAAAACTCCCCC-----AGACCGAGTTGGG--CCAACCTT 900
Db 516 AGAGGAAAAACAGAAAGTTGAAAAAACTCCCCC-----AGAGTGGGGCCCAACTTT 575
QY 901 TGATCAGAGTCC--AGGACGAGTCGAGGTGGCTGGCTCTCTTTTGGGCGCGGTCTGGAA 958
Db 576 GGATCAGAGTCCCGAGGACCCAGTGCAGGTGGCTGGCTCTCTTTTGGGCGCGGTCTGGAA 635
QY 959 TAATGGACGCGCTGGCAGTCCAGACATCAACTCCAAAACTGAAGCTGCAGCAATGAAGA 1018
Db 636 AATCGACGCGCTGGCAGTCCAGACATCAATCAAACTGAAGCTGCAGCAATGAAGA 695
QY 1019 AGCAGTCACATACAGAAAAAGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAG 1078
Db 696 AGCAGTCACATACAGAAAAAGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAG 755
QY 1079 CAAAAGTCAACAAACCCCTATTATACCTTCACCCAAAATTTTATCATTTGCTCTTTCTT 1138
Db 756 C--AAAGTCAAAACCCCTATTATACCTTCCA-CCAAATTTCTTTATCATTTGCTCTTCT 812
QY 1139 AGGAAACAGACATCTCATTTGATTTTAAATAAGTTTATTTT 1185
Db 813 AGGAAACAGACATCTCATTTGATTTTAAATAAGTTTAAATTT 859

RESULT 8

US-10-450-763-19570
; Sequence 19570, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 19570
; LENGTH: 742
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (307)..(642)
; OTHER INFORMATION: 67% homologous to Homo sapiens TRAF2 binding protein encoded by clone 15, accession number W42401, Smith-Waterman Score=349.
US-10-450-763-19570

Query Match 29.6%; Score 370.4; DB 8; Length 742;
Best Local Similarity 97.4%; Pred. No. 3.5e-86;
Matches 419; Conservative 0; Mismatches 6; Indels 5; Gaps 4;

QY 757 AGGTGCCACACCTCCCTGGATGATCCAGATGAAGATACATTTCTGGGACCAAGAAAT 816
Db 306 AGGTGCCACACCTCCCTGGATGATCCAGATGAAGATACATTTCTGGGACCAAGAAAT 365

Db 185 ACAGCCTGGAGTCCAGACATCAA-TTCAAACTGAAGCTGCAGCAATGAAGAAGCAGT 127
QY 1025 CACATACAGAAAAAGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCCAAAG 1084
Db 126 CACATACAGAAAAAGCTAATCATG--CTTTACCAACTACCATGAGGCTAAAAG-CAAAG 70
QY 1085 TCAACCAAAACCCCTATTATACCTTCCACCCTTCTTTATCATGCTCTTTCTTAGGAAA 1144
Db 69 TCAACAAACCCC-----TACTTTCCACCGAATCTTTATCATGCTCTTTCTTAGGAAA 17
QY 1145 CAGACATACCTATTCA 1160
Db 16 CAGACATACCTATTCA 1

RESULT 11

US-60-680-473-781
; Sequence 781, Application US/60680473
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary to
; FILE REFERENCE: 21590290000
; CURRENT APPLICATION NUMBER: US/60/680,473
; NUMBER OF SEQ ID NOS: 2005-05-13
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 781
; LENGTH: 743
; TYPE: DNA
; ORGANISM: Macaca Fascicularis
; NAME/KEY: misc feature
; LOCATION: (1)...(743)
; OTHER INFORMATION: n = A,T,C or G
US-60-680-473-781

Query Match 25.8%; Score 323.6; DB 15; Length 743;
Best Local Similarity 93.3%; Pred. No. 5.5e-74;
Matches 407; Conservative 0; Mismatches 14; Indels 15; Gaps 6;
QY 725 CAGGATATACAGGAGTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAA 784
Db 323 CAGGATATACAGGAGTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAA 382
QY 785 GATGAAGATACATTGCTGGGAACCAAGAAATAGACCATCTTATGAAGAATTTCTTAAA 844
Db 383 GAT---GAATACATCGTGGGAACCAAGAAATAGACCATCTTATGAAGAATTTCTTAAA 439
QY 845 GAAAGGAAAAACAGAAGTTGAAAAAACTCCCCCAGACCGAGTTGGGCCCACTTTGAT 904
Db 440 GAAAGGAAAAACAGAAGTTGAAAAAACTCCCCCAGACCGAGTTGGGCCCACTTTGAT 499
QY 905 CACAGCTCCAGGACCAAGTTCGAGCTGGCTGGCTCTCTTTTGGGCCCGCTTGGAAATAATGG 964
Db 500 CACAGCTCCAGGACCAAGTTCGAGCTGGCTGGCTCTCTTTTGGGCCCGCTTGGAAATAATGG 558
QY 965 ACGCCGCTGGCAGTCCAGACATCACTCCAAAACCTGAAGCTGCAGCATGAAGAAGCAGT 1024
Db 559 ACAGCGCTGGCAGTCCAGACATCAA-TTCAAACTGAAGCTGCAGCATGAAGAAGCAGT 617
QY 1025 CACATACAGAAAAAGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCCAAAG 1084
Db 618 CACATACAGAAAAAGCTAATCATG--CTTTACCAACTACCATGAGGCTAAAAG-CAAAG 674
QY 1085 TCAACCAAAACCCCTATTATACCTTCCACCCTTCTTTATCATGCTCTTTCTTAGGAAA 1144

Db 675 TCAACAAACCCC-----TACTTTCCACCGAATCTTTATCATGCTCTTTCTTAGGAAA 727
QY 1145 CAGACATACCTATTCA 1160
Db 728 CAGACATACCTATTCA 743
RESULT 12
US-60-680-473-7415/c
; Sequence 7415, Application US/60680473
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary to
; FILE REFERENCE: 21590290000
; CURRENT APPLICATION NUMBER: US/60/680,473
; NUMBER OF SEQ ID NOS: 2005-05-13
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 7415
; LENGTH: 743
; TYPE: DNA
; ORGANISM: Macaca Fascicularis
; NAME/KEY: misc feature
; LOCATION: (1)...(743)
; OTHER INFORMATION: n = A,T,C or G
US-60-680-473-7415

Query Match 25.8%; Score 323.6; DB 15; Length 743;
Best Local Similarity 93.3%; Pred. No. 5.5e-74;
Matches 407; Conservative 0; Mismatches 14; Indels 15; Gaps 6;
QY 725 CAGGATATACAGGAGTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAA 784
Db 421 CAGGATATACAGGAGTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAA 362
QY 785 GATGAAGATACATTGCTGGGAACCAAGAAATAGACCATCTTATGAAGAATTTCTTAAA 844
Db 361 GAT---GAATACATCGTGGGAACCAAGAAATAGACCATCTTATGAAGAATTTCTTAAA 305
QY 845 GAAAGGAAAAACAGAAGTTGAAAAAACTCCCCCAGACCGAGTTGGGCCCACTTTGAT 904
Db 304 GAAAGGAAAAACAGAAGTTGAAAAAACTCCCCCAGACCGAGTTGGGCCCACTTTGAT 245
QY 905 CACAGCTCCAGGACCAAGTTCGAGCTGGCTGGCTCTCTTTTGGGCCCGCTTGGAAATAATGG 964
Db 244 CACAGCTCCAGGACCAAGTTCGAGCTGGCTGGCTCTCTTTTGGGCCCGCTTGGAAATAATGG 186
QY 965 ACGCCGCTGGCAGTCCAGACATCACTCCAAAACCTGAAGCTGCAGCATGAAGAAGCAGT 1024
Db 185 ACAGCGCTGGCAGTCCAGACATCAA-TTCAAACTGAAGCTGCAGCATGAAGAAGCAGT 127
QY 1025 CACATACAGAAAAAGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCCAAAG 1084
Db 126 CACATACAGAAAAAGCTAATCATG--CTTTACCAACTACCATGAGGCTAAAAG-CAAAG 70
QY 1085 TCAACCAAAACCCCTATTATACCTTCCACCCTTCTTTATCATGCTCTTTCTTAGGAAA 1144
Db 69 TCAACAAACCCC-----TACTTTCCACCGAATCTTTATCATGCTCTTTCTTAGGAAA 17
QY 1145 CAGACATACCTATTCA 1160
Db 16 CAGACATACCTATTCA 1

Mon Jun 13 13:27:14 2005

Db 808 GGTGCCACACCTCCCTGGATGATCCAGATGAAGATACATTGCTGGAAACCAAGAATA 867
Qy 818 GGACCATCCTATGAAGAATTTCCTTAAGAAAAGGAAAAACAGAAAGTTGAAAAAACTCCCC 877
Db 868 GGACCATCCTATGAAGAATTTCCTTAAGAAAAGGAAAAACAGAAAGTTGAAAAAACTCCCC 927
Qy 878 CCAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCAAGTGCAGGCTGGCTGCC 937
Db 928 CCAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCAAGTGCAGGCTGGCTGCC 987
Qy 938 TCTTTTGGGCGCGTCTGGAATAATGGACCGCGCTGGCAGTCCAGACATCAA 989
Db 988 TCTTTT-GGCCGCGTCTGGAATAATGGACCGCGCTGGCAGTCCAGAGCTCCA 1038

Search completed: June 11, 2005, 05:07:14
Job time : 482.315 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 06:13:19 ; Search time 3133.95 Seconds
(without alignments)
16321.674 Million cell updates/sec

Title: US-09-155-676B-4

Perfect score: 1253

Sequence: 1 cattggagtcacgcggtggc.....ggggggcgctaccacatttt 1253

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA Main:

- 1: /cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/PCTUS3_COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US08_COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
- 8: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
- 9: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
- 10: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
- 12: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
- 13: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
- 14: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
- 15: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
- 16: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
- 17: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
- 18: /cgn2_6/ptodata/1/pna/US092A_COMB.seq.*
- 19: /cgn2_6/ptodata/1/pna/US092B_COMB.seq.*
- 20: /cgn2_6/ptodata/1/pna/US093A_COMB.seq.*
- 21: /cgn2_6/ptodata/1/pna/US093B_COMB.seq.*
- 22: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
- 23: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
- 24: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
- 25: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
- 26: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
- 27: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
- 28: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
- 29: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
- 30: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
- 31: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
- 32: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
- 33: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
- 34: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
- 35: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
- 36: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
- 37: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
- 38: /cgn2_6/ptodata/1/pna/US098D_COMB.seq.*
- 39: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
- 40: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
- 41: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
- 42: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
- 43: /cgn2_6/ptodata/1/pna/US099E_COMB.seq.*

- 44: /cgn2_6/ptodata/1/pna/US099F_COMB.seq.*
- 45: /cgn2_6/ptodata/1/pna/US099G_COMB.seq.*
- 46: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
- 47: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
- 48: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
- 49: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
- 50: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
- 51: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*
- 52: /cgn2_6/ptodata/1/pna/US103A_COMB.seq.*
- 53: /cgn2_6/ptodata/1/pna/US103B_COMB.seq.*
- 54: /cgn2_6/ptodata/1/pna/US104A_COMB.seq.*
- 55: /cgn2_6/ptodata/1/pna/US104B_COMB.seq.*
- 56: /cgn2_6/ptodata/1/pna/US105A_COMB.seq.*
- 57: /cgn2_6/ptodata/1/pna/US105B_COMB.seq.*
- 58: /cgn2_6/ptodata/1/pna/US106A_COMB.seq.*
- 59: /cgn2_6/ptodata/1/pna/US107A_COMB.seq.*
- 60: /cgn2_6/ptodata/1/pna/US107B_COMB.seq.*
- 61: /cgn2_6/ptodata/1/pna/US107C_COMB.seq.*
- 62: /cgn2_6/ptodata/1/pna/US107D_COMB.seq.*
- 63: /cgn2_6/ptodata/1/pna/US108A_COMB.seq.*
- 64: /cgn2_6/ptodata/1/pna/US108B_COMB.seq.*
- 65: /cgn2_6/ptodata/1/pna/US109A_COMB.seq.*
- 66: /cgn2_6/ptodata/1/pna/US109B_COMB.seq.*
- 67: /cgn2_6/ptodata/1/pna/US109C_COMB.seq.*
- 68: /cgn2_6/ptodata/1/pna/US110_COMB.seq.*
- 69: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
- 70: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
- 71: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
- 72: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
- 73: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
- 74: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
- 75: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
- 76: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
- 77: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
- 78: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
- 79: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
- 80: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
- 81: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
- 82: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
- 83: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
- 84: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
- 85: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
- 86: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
- 87: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
- 88: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
- 89: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
- 90: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
- 91: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
- 92: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq.*
- 93: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq.*
- 94: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
- 95: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
- 96: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
- 97: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
- 98: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
- 99: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
- 100: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
- 101: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
- 102: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
- 103: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
- 104: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
- 105: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
- 106: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
- 107: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
- 108: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
- 109: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
- 110: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*
- 111: /cgn2_6/ptodata/1/pna/US6042_COMB.seq.*
- 112: /cgn2_6/ptodata/1/pna/US6043_COMB.seq.*
- 113: /cgn2_6/ptodata/1/pna/US6044_COMB.seq.*
- 114: /cgn2_6/ptodata/1/pna/US6045_COMB.seq.*
- 115: /cgn2_6/ptodata/1/pna/US6046_COMB.seq.*
- 116: /cgn2_6/ptodata/1/pna/US6046_COMB.seq.*

117: /cgn2_6/ptodata/1/pna/US6047_COMB.seq:*
118: /cgn2_6/ptodata/1/pna/US6048_COMB.seq:*
119: /cgn2_6/ptodata/1/pna/US6049_COMB.seq:*
120: /cgn2_6/ptodata/1/pna/US6050_COMB.seq:*
121: /cgn2_6/ptodata/1/pna/US6051_COMB.seq:*
122: /cgn2_6/ptodata/1/pna/US6052_COMB.seq:*
123: /cgn2_6/ptodata/1/pna/US6053_COMB.seq:*
124: /cgn2_6/ptodata/1/pna/US6054_COMB.seq:*
125: /cgn2_6/ptodata/1/pna/US6055_COMB.seq:*
126: /cgn2_6/ptodata/1/pna/US6056_COMB.seq:*
127: /cgn2_6/ptodata/1/pna/US6057_COMB.seq:*
128: /cgn2_6/ptodata/1/pna/US6058_COMB.seq:*
129: /cgn2_6/ptodata/1/pna/US6059_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1252	99.9	1253	17	US-09-155-676A-4
2	1252	99.9	1253	17	US-09-155-676A-4
3	1252	99.9	1253	124	US-60-545-213-1553
4	1252	99.9	1253	124	US-60-545-213-1554
5	1252	99.9	1253	124	US-60-545-213-15825
6	1252	99.9	1253	124	US-60-545-213-5826
7	1021.8	81.5	1200	22	US-09-471-275-1856
8	1021.8	81.5	1200	22	US-09-488-725B-6991
9	1021.8	81.5	1200	26	US-09-552-317-6991
10	997.4	79.6	1355	33	US-09-758-463-382
11	997.4	79.6	1355	50	US-10-217-607-382
12	663.4	52.9	1415	90	US-60-213-362-7088
13	663.4	52.9	3215	90	US-60-213-362-1179
14	663.4	52.9	3216	97	US-60-278-561-2059
15	661.8	52.8	2561	102	US-60-324-185-7629
16	656	52.4	1191	97	US-60-278-561-3149
17	648.2	51.7	1637	31	US-09-698-012-8348
18	648.2	51.7	1637	32	US-09-716-920-1616
19	644.4	51.4	1262	18	US-09-205-070-15304
20	644.4	51.4	1262	20	US-09-340-623-15304
21	644.4	51.4	1262	38	US-09-898-888-15304
22	644.4	51.4	1262	38	US-09-898-888A-15304
23	634.8	50.7	3252	97	US-60-278-232-9368
24	595	47.5	1318	21	US-09-359-922-12955
25	595	47.5	1318	21	US-09-359-922-12955
26	595	47.5	1318	39	US-09-919-002-12955
27	580.8	46.4	2273	86	US-60-172-373-23318
28	548.6	43.8	881	92	US-60-234-690-3752
29	548.6	43.8	881	102	US-60-323-843-3752
30	514.6	41.1	681	1	PCT-US01-08631-19571
31	487.4	38.9	577	24	US-09-539-331D-31548
32	487	38.9	764	30	US-09-668-683-14062
33	479.6	38.3	2557	1	PCT-US01-08631-23046
34	451.8	36.1	687	49	US-10-170-235-38959
35	451.8	36.1	3948	39	US-09-915-543-16
36	451.8	36.1	3948	52	US-10-322-579-16
37	451.8	36.1	3948	58	US-10-664-859-16
38	451.4	36.0	687	115	US-60-453-050-1425
39	451.4	36.0	687	115	US-60-453-135-1425
40	451.4	36.0	687	116	US-60-466-412-1425
41	444.4	35.5	500	47	US-10-085-783A-45728
42	444.4	35.5	500	50	US-10-242-535A-45728
43	428.2	34.2	444	21	US-09-399-720-4897
44	428.2	34.2	444	40	US-09-921-378-4897
45	407	32.5	430	20	US-09-321-214-15129

ALIGNMENTS

RESULT 1
US-09-155-676-4
; Sequence 4: Application US/09155676
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/155,676
; FILING DATE: 04-JAN-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IL97/00117
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117800
; FILING DATE: 02-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 119133
; FILING DATE: 26-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: WALLACH=21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1253 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-155-676-4
Query Match 99.9%; Score 1252; DB 17; Length 1253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATTGGAGTCACGGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60
Db 1 CATTGGAGTCACGGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60
QY 61 GATTTCAGAGCCCAAGAGGCCCTTCTTCTGTGGTGGCGCAGTTTACAGCCGCAAGCAC 120
Db 61 GATTTCAGAGCCCAAGAGGCCCTTCTTCTGTGGTGGCGCAGTTTACAGCCGCAAGCAC 120
QY 121 CCAGCGGCAAGTCAGAGGAGGCTTTTGGAGAGGCTCTCTGCCCCAGGTGGAGCGCGCCCGCAA 180
Db 121 CCAGCGGCAAGTCAGAGGAGGCTTTTGGAGAGGCTCTCTGCCCCAGGTGGAGCGCGCCCGCAA 180
QY 181 GGCATCCCGCGCGCTCAGGTGGAGCGCTATGTGCCCGCAACACGAGCGATGCTGCTGGTG 240
Db 181 GGCATCCCGCGCGCTCAGGTGGAGCGCTATGTGCCCGCAACACGAGCGATGCTGCTGGTG 240

Qy	241	CCTGTGCTGGCGCTGTGAGGTGCGGGAAACACCTGAGCCATGGAAACCTGACGGTCTGTGA	300
Db	241	CCTGTGCTGGCGCTGTGAGGTGCGGGAAACACCTGAGCCATGGAAACCTGACGGTCTGTGA	300
Qy	301	CGGGGGCTGTGTGAGCATCTGGCCAGCCAGAGCACAAAGAAAGCAACCAACAATTTCTG	360
Db	301	CGGGGGCTGTGTGAGCATCTGGCCAGCCAGAGCACAAAGAAAGCAACCAACAATTTCTG	360
Qy	361	GTGGGAGAACAAAGCTGAGTCTCAGATGAAAGAGAAAGTTTCTGTCTCATCCCCAGGATTA	420
Db	361	GTGGGAGAACAAAGCTGAGTCTCAGATGAAAGAGAAAGTTTCTGTCTCATCCCCAGGATTA	420
Qy	421	TGCGCGATTCAAGAAATCCATGSGTGAAGGTTTGGATTCTTATCAAGAAAGAGGATAA	480
Db	421	TGCGCGATTCAAGAAATCCATGSGTGAAGGTTTGGATTCTTATCAAGAAAGAGGATAA	480
Qy	481	AGTGATCAAGGAGATGGCAGCTCAGATCCCGTGAGGTGGAGCAGAGCCGACAGAGAGTGGT	540
Db	481	AGTGATCAAGGAGATGGCAGCTCAGATCCCGTGAGGTGGAGCAGAGCCGACAGAGAGTGGT	540
Qy	541	TCGGTCTGTCTTAGAGCTTCAGGCATGCGCAGACCCAGAAAGAGGGCTCTTCAGCACCTAG	600
Db	541	TCGGTCTGTCTTAGAGCTTCAGGCATGCGCAGACCCAGAAAGAGGGCTCTTCAGCACCTAG	600
Qy	601	AAGCTGAAAGGGATGAACACGCCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAATTTGGA	660
Db	601	AAGCTGAAAGGGATGAACACGCCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAATTTGGA	660
Qy	661	CCTGCCACCACTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGG	720
Db	661	CCTGCCACCACTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGG	720
Qy	721	CCATCAGGATATACCGAGGTTGGTAAATCATCCAATCAGGTGCCACACTCTCCCTGGATGAT	780
Db	721	CCATCAGGATATACCGAGGTTGGTAAATCATCCAATCAGGTGCCACACTCTCCCTGGATGAT	780
Qy	781	CCAAGATGAAGAATACATCTGCTGGAAACCAAGAAATAGGACCATCTTATGAAGAATTTCT	840
Db	781	CCAAGATGAAGAATACATCTGCTGGAAACCAAGAAATAGGACCATCTTATGAAGAATTTCT	840
Qy	841	TAAAGAAAAGGAAAAACAGAAAGTTGAAAAAATCTCCCCCAGACCGAGTTGGGGCCAACTT	900
Db	841	TAAAGAAAAGGAAAAACAGAAAGTTGAAAAAATCTCCCCCAGACCGAGTTGGGGCCAACTT	900
Qy	901	TGATCAGCTCCAGGACCACTGAGGCTGGCTGCCCTCTTTTGGGCGGGCTCTGGAAATA	960
Db	901	TGATCAGCTCCAGGACCACTGAGGCTGGCTGCCCTCTTTTGGGCGGGCTCTGGAAATA	960
Qy	961	ATGGACGCGCTGGCAGTCCAGACATCAACTCCAAAACTGAAGCTGCAGCGAATGAAGAAG	1020
Db	961	ATGGACGCGCTGGCAGTCCAGACATCAACTCCAAAACTGAAGCTGCAGCGAATGAAGAAG	1020
Qy	1021	CAGTCAATACAGAAAAAAGCTAATCATGTCTCTCAACAACTACCATGAGGCTAAAAACC	1080
Db	1021	CAGTCAATACAGAAAAAAGCTAATCATGTCTCTCAACAACTACCATGAGGCTAAAAACC	1080
Qy	1081	AAAGTCAACAAACCCCTATTATACCTTCGACCCAAATTCCTTTATCATGTCTTTCTTAG	1140
Db	1081	AAAGTCAACAAACCCCTATTATACCTTCGACCCAAATTCCTTTATCATGTCTTTCTTAG	1140
Qy	1141	GAAACAGACATACCTCATTTGATTTAATAAAGTTTATTTTTCGGCCCTTCGTGGCCT	1200
Db	1141	GAAACAGACATACCTCATTTGATTTAATAAAGTTTATTTTTCGGCCCTTCGTGGCCT	1200
Qy	1201	CGAATCAAGCTTATCGATACCGTCCGACTCGAGGGGGGGCGGTACCACTTTT	1253
Db	1201	CGAATCAAGCTTATCGATACCGTCCGACTCGAGGGGGGGCGGTACCACTTTT	1253

RESULT 2

RESULT 2
US-09-155-676A-4

US-09-133-676A-4
; Sequence 4, Application US/09155676A; sequence 4, application:
; GENERAL INFORMATION:

APPLICANT: WALLACH, David
APPLICANT: MALININ, Nikolai
APPLICANT: BOLDIN, Mark
APPLICANT: KOVALENKO, Andrei
APPLICANT: METT, Igor
TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
STREET: 624 Ninth Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/155,676A
FILING DATE: 04-JAN-1999
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IL97/00117
FILING DATE: 01-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 117800
FILING DATE: 02-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 119133
FILING DATE: 26-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH=21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1253 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
PS-09-155-676A-4

Query Match	99.9%	Score 1252;	DB 17;	Length 1253;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1253:	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

Qy	1	CATTGGAGTCA	CGGGTGGCGGCTCTAGAA	TAGTGGATCCC	CGGGCTGC	GANGAAATTC	60
Db	1	CATTGGAGTCA	CGGGTGGCGGCTCTAGAA	TAGTGGATCCC	CGGGCTGC	GANGAAATTC	60
Qy	61	GATTTCGAGCC	ACGAAGGCCCTCTTCTCTGTGTGGCGG	CACGTTTAC	GAGCGCAAGCAC	120	
Db	61	GATTTCGAGCC	ACGAAGGCCCTCTTCTCTGTGTGGCGG	CACGTTTAC	GAGCGCGAAGCAC	120	
Qy	121	CCAGCGCAGCT	GAAAGAGGCTTTTGAGAGGCTCTTGCC	CAGGTGAGGCGGCGCCGCA	180		
Db	121	CCAGCGCAGCT	GAAAGAGGCTTTTGAGAGGCTCTTGCC	CAGGTGAGGCGGCGCCGCA	180		
Qy	181	GGCCATCCGCG	CCGCTCAGGTGGAGCGCTATGTGCCGGA	ACACGAGCGATGCTCTGCTGTG	240		
Db	181	GGCCATCCGCG	CCGCTCAGGTGGAGCGCTATGTGCCGGA	ACACGAGCGATGCTCTGCTGTG	240		
Qy	241	CCTGTGTCGGG	CTGTGAGTGGCGGGAACACCTGAGCCA	TGGAAACCTGACGGTGCTGTGA	300		
Db	241	CCTGTGTCGGG	CTGTGAGTGGCGGGAACACCTGAGCCA	TGGAAACCTGACGGTGCTGTGA	300		

```
QY 301 CGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCACAAAGAAAGCAACCAAAATTCG 360
Db |||||||
QY 301 CGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCACAAAGAAAGCAACCAAAATTCG 360
Db |||||||
QY 361 GTGGGAGAACAAAGCTGAGGTCAGATGAAAGAGAGTTCTTGTCATCTCCCAAGATTA 420
Db |||||||
QY 361 GTGGGAGAACAAAGCTGAGGTCAGATGAAAGAGAGTTCTTGTCATCTCCCAAGATTA 420
Db |||||||
QY 421 TCGCGGATCAAGAAATCCATGGTGAAGTTGATTCCTATGAAGAAAGGAGGATAA 480
Db |||||||
QY 421 TCGCGGATCAAGAAATCCATGGTGAAGTTGATTCCTATGAAGAAAGGAGGATAA 480
Db |||||||
QY 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGGT 540
Db |||||||
QY 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGGT 540
Db |||||||
QY 541 TCGGTCCTGTCTTAGAGCTTCAGGCTCAGAGCCAGAGAGGCTCTTCAGACCTAG 600
Db |||||||
QY 541 TCGGTCCTGTCTTAGAGCTTCAGGCTCAGAGCCAGAGAGGCTCTTCAGACCTAG 600
Db |||||||
QY 601 AAGCTGGAAGGATGAACAGCTGAGGTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGGT 660
Db |||||||
QY 601 AAGCTGGAAGGATGAACAGCTGAGGTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGGT 660
Db |||||||
QY 661 CTGCGCAGCTCAGAGCTTCAGGCTCAGAGCCAGAGAGGCTCTTCAGACCTAG 720
Db |||||||
QY 661 CTGCGCAGCTCAGAGCTTCAGGCTCAGAGCCAGAGAGGCTCTTCAGACCTAG 720
Db |||||||
QY 721 CCATCAGGATATACCAAGAGTTGGTAACTCACTCAGTGGAGCAGAGCCATCTCTGACATTCATTGG 780
Db |||||||
QY 721 CCATCAGGATATACCAAGAGTTGGTAACTCACTCAGTGGAGCAGAGCCATCTCTGACATTCATTGG 780
Db |||||||
QY 781 CCAAGATGAAGATACATCTGCTGGAGCAAGAAATAGGACCATCTATGAAGATTTCT 840
Db |||||||
QY 781 CCAAGATGAAGATACATCTGCTGGAGCAAGAAATAGGACCATCTATGAAGATTTCT 840
Db |||||||
QY 841 TAAAGAAAGGAAAAACAGAGTTGAAAAAATCTCCCGCCAGAGCCGAGTTGGGGCCAACTT 900
Db |||||||
QY 841 TAAAGAAAGGAAAAACAGAGTTGAAAAAATCTCCCGCCAGAGCCGAGTTGGGGCCAACTT 900
Db |||||||
QY 901 TGATCAGCTCAGAGCAGTGCAGGCTGGCTGGCTCTTTTGGCCCGCTCTGGAATA 960
Db |||||||
QY 901 TGATCAGCTCAGAGCAGTGCAGGCTGGCTGGCTCTTTTGGCCCGCTCTGGAATA 960
Db |||||||
QY 961 ATGAGCGCTGCGAGTCCAGACATCACTCCAAACTGAAGCTCAGCAATGAAGAAG 1020
Db |||||||
QY 961 ATGAGCGCTGCGAGTCCAGACATCACTCCAAACTGAAGCTCAGCAATGAAGAAG 1020
Db |||||||
QY 1021 CAGTCACATACAGAAAAAGCTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCC 1080
Db |||||||
QY 1021 CAGTCACATACAGAAAAAGCTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCC 1080
Db |||||||
QY 1081 AAGTCAACCAACCCCTATTATACCTCCACCAAAATCTTTATCATTTCTTTCTTAG 1140
Db |||||||
QY 1081 AAGTCAACCAACCCCTATTATACCTCCACCAAAATCTTTATCATTTCTTTCTTAG 1140
Db |||||||
QY 1141 GAAACAGACATCTCATTTGATTTAATAAGTTTATTTTTCGGCCCTTCGTGGCCT 1200
Db |||||||
QY 1141 GAAACAGACATCTCATTTGATTTAATAAGTTTATTTTTCGGCCCTTCGTGGCCT 1200
Db |||||||
QY 1201 CGAATCAAGCTTATCGATACCGTCCGACCTCGAGGGGGGGCGGACCACTTTT 1253
Db |||||||
QY 1201 CGAATCAAGCTTATCGATACCGTCCGACCTCGAGGGGGGGCGGACCACTTTT 1253
Db |||||||
```

RESULT 3
US-60-545-213-1553
; Sequence 1553, Application US/60545213
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes

FILE REFERENCE: AM101083 (031896-042099)
CURRENT APPLICATION NUMBER: US/60/545,213
CURRENT FILING DATE: 2004-02-18
NUMBER OF SEQ ID NOS: 303284
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1553
LENGTH: 1253
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (53)..(53)
OTHER INFORMATION: n is a, c, g, or t
US-60-545-213-1553

Query Match		99,9%	Score 1252;	DB 124;	Length 1253;
Best Local Similarity		100.0%	Pred. No. 0;		
Matches 1253;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CATTGGAGTCAACGCGGTGGCGGCTCTAGAAATAGTGATCCCGGGGCTGCAGGAATTC	60		
Db	1	CATTGGAGTCAACGCGGTGGCGGCTCTAGAAATAGTGATCCCGGGGCTGCAGGAATTC	60		
QY	61	GATTGAGCCCAAGAGGCGGCTTTCTGTGTGCGGCGACGTTTACAGCCGCAAGCAC	120		
Db	61	GATTGAGCCCAAGAGGCGGCTTTCTGTGTGCGGCGACGTTTACAGCCGCAAGCAC	120		
QY	121	CCAGCGGCGAGTGAAGAGAGGCTTTTGAGAGGCTCTGCCCCAGGTGGAGGCGGCGCA	180		
Db	121	CCAGCGGCGAGTGAAGAGAGGCTTTTGAGAGGCTCTGCCCCAGGTGGAGGCGGCGCA	180		
QY	181	GGCCATCCGCGCGCTCAGGTGGAGCGCTATGTGCCCGAACAAGAGCGATCTGCTGTG	240		
Db	181	GGCCATCCGCGCGCTCAGGTGGAGCGCTATGTGCCCGAACAAGAGCGATCTGCTGTG	240		
QY	241	CCTGTGCTCGCGCTGTGAGGTGCGGGAACCTTGAGCCATGAAACCTGACGCTGCTGA	300		
Db	241	CCTGTGCTCGCGCTGTGAGGTGCGGGAACCTTGAGCCATGAAACCTGACGCTGCTGA	300		
QY	301	CGGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCAAGAAAGCAACAAATTCG	360		
Db	301	CGGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCAAGAAAGCAACAAATTCG	360		
QY	361	GTGGAGAACAAAGCTGAGGTCCAGATGAAGAGAGTTCTTGTCCTCCCAAGATTA	420		
Db	361	GTGGAGAACAAAGCTGAGGTCCAGATGAAGAGAGTTCTTGTCCTCCCAAGATTA	420		
QY	421	TGCGGATTCAGAAATCCATGGTGAAGTTTGGATTCCTATGAAGAAAGGAGATTA	480		
Db	421	TGCGGATTCAGAAATCCATGGTGAAGTTTGGATTCCTATGAAGAAAGGAGATTA	480		
QY	481	AGTGATCAAGGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGGAGTGT	540		
Db	481	AGTGATCAAGGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGGAGTGT	540		
QY	541	TCGCTCTGTCTTAGAGCTCAGGAGTCCAGAGGAGGCTCTTCAGACCTAG	600		
Db	541	TCGCTCTGTCTTAGAGCTCAGGAGTCCAGAGGAGGCTCTTCAGACCTAG	600		
QY	601	AAGCTGGAAGGATGAACAGCTGAGTCCAGAGTTCAGCTTACAGAGCCCTCAAAATTTGA	660		
Db	601	AAGCTGGAAGGATGAACAGCTGAGTCCAGAGTTCAGCTTACAGAGCCCTCAAAATTTGA	660		
QY	661	CTTGCCACAGCTCCAGAGTTCAGTGGATGGAGCAGGACCATCTCTGACATTCATTGG	720		
Db	661	CTTGCCACAGCTCCAGAGTTCAGTGGATGGAGCAGGACCATCTCTGACATTCATTGG	720		
QY	721	CCATCAGGATATACCAAGAGTTGGTAACTCACTCAGTGGAGCAGAGCCATCTCTGACATTCATTGG	780		
Db	721	CCATCAGGATATACCAAGAGTTGGTAACTCACTCAGTGGAGCAGAGCCATCTCTGACATTCATTGG	780		
QY	781	CCAAGATGAAGATACATCTGCTGGAGCAAGAAATAGGACCATCTATGAAGATTTCT	840		
Db	781	CCAAGATGAAGATACATCTGCTGGAGCAAGAAATAGGACCATCTATGAAGATTTCT	840		
QY	841	TAAAGAAAGGAAAAACAGAGTTGAAAAAATCTCCCGCCAGAGCCGAGTTGGGGCCAACTT	900		
Db	841	TAAAGAAAGGAAAAACAGAGTTGAAAAAATCTCCCGCCAGAGCCGAGTTGGGGCCAACTT	900		
QY	901	TGATCAGCTCAGAGCAGTGCAGGCTGGCTGGCTCTTTTGGCCCGCTCTGGAATA	960		
Db	901	TGATCAGCTCAGAGCAGTGCAGGCTGGCTGGCTCTTTTGGCCCGCTCTGGAATA	960		
QY	961	ATGAGCGCTGCGAGTCCAGACATCACTCCAAACTGAAGCTCAGCAATGAAGAAG	1020		
Db	961	ATGAGCGCTGCGAGTCCAGACATCACTCCAAACTGAAGCTCAGCAATGAAGAAG	1020		
QY	1021	CAGTCACATACAGAAAAAGCTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCC	1080		
Db	1021	CAGTCACATACAGAAAAAGCTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCC	1080		
QY	1081	AAGTCAACCAACCCCTATTATACCTCCACCAAAATCTTTATCATTTCTTTCTTAG	1140		
Db	1081	AAGTCAACCAACCCCTATTATACCTCCACCAAAATCTTTATCATTTCTTTCTTAG	1140		
QY	1141	GAAACAGACATCTCATTTGATTTAATAAGTTTATTTTTCGGCCCTTCGTGGCCT	1200		
Db	1141	GAAACAGACATCTCATTTGATTTAATAAGTTTATTTTTCGGCCCTTCGTGGCCT	1200		
QY	1201	CGAATCAAGCTTATCGATACCGTCCGACCTCGAGGGGGGGCGGACCACTTTT	1253		
Db	1201	CGAATCAAGCTTATCGATACCGTCCGACCTCGAGGGGGGGCGGACCACTTTT	1253		

Db 781 CCAAGATGAAGATACATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGATTCT 840
Qy 841 TAAAGAAAGGAAAGAAACAGAGATTGAAAGAACTCCCCCCAGACGAGTTGGGGCAACTT 900
Db 841 TAAAGAAAGGAAAGAAACAGAGATTGAAAGAACTCCCCCCAGACGAGTTGGGGCAACTT 900
Qy 901 TGATCACAGCTCCAGGACCAAGTGCAGGCTGCTGCTCTCTTTTGGGCGGCTCTGGAATA 960
Db 901 TGATCACAGCTCCAGGACCAAGTGCAGGCTGCTGCTCTCTTTTGGGCGGCTCTGGAATA 960
Qy 961 ATGGAGCGCGCTGGCAGTCCAGACATCAATCCAAAACCTGAAGTGCAGCAATGAAGAAG 1020
Db 961 ATGGAGCGCGCTGGCAGTCCAGACATCAATCCAAAACCTGAAGTGCAGCAATGAAGAAG 1020
Qy 1021 CAGTCACATACAGAAAAAGCTAATCATGTCTCTCAACAATCAATGAGGCTAAAAGCC 1080
Db 1021 CAGTCACATACAGAAAAAGCTAATCATGTCTCTCAACAATCAATGAGGCTAAAAGCC 1080
Qy 1081 AAAGTCAACCAACCCCTAATATACCTTCCACCCAAATCTTTATCATGTTCTTCTTAG 1140
Db 1081 AAAGTCAACCAACCCCTAATATACCTTCCACCCAAATCTTTATCATGTTCTTCTTAG 1140
Qy 1141 GAAACAGACATCACTCATTCATTGATTTAATAAAGTTTATTTTCGGCCTTCGTGGCCT 1200
Db 1141 GAAACAGACATCACTCATTCATTGATTTAATAAAGTTTATTTTCGGCCTTCGTGGCCT 1200
Qy 1201 CGAATCAAGCTTATCGATACCGTCCAGCTCGAGCTGGGGGGGGGCTACCCACTTTT 1253
Db 1201 CGAATCAAGCTTATCGATACCGTCCAGCTCGAGCTGGGGGGGGGCTACCCACTTTT 1253

RESULT 4

US-60-545-213-1554
; Sequence 1554, Application US/60545213
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042099)
; CURRENT APPLICATION NUMBER: US/60/545,213
; CURRENT FILING DATE: 2004-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1554
; LENGTH: 1253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: n is a, c, g, or t
US-60-545-213-1554

Query Match 99.9%; Score 1252; DB 124; Length 1253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CATTTGGAGTCACGCGTGGCGGCTCTAGATAGTATGATCCCGGGCTGCAGGAATTC 60
Db 1 CATTTGGAGTCACGCGTGGCGGCTCTAGATAGTATGATCCCGGGCTGCAGGAATTC 60
Qy 61 GATTTCGAGCCCAAGAGCCCTTCTTCTGTGTGTGGGACAGTTTACAGCCGCAAGCAC 120
Db 61 GATTTCGAGCCCAAGAGCCCTTCTTCTGTGTGTGGGACAGTTTACAGCCGCAAGCAC 120
Qy 121 CCAGCGGAGCTGAAGAGGCTTTTGAGAGCTCTGCCCCAGGTGAGGCGGCCGCCAA 180
Db 121 CCAGCGGAGCTGAAGAGGCTTTTGAGAGCTCTGCCCCAGGTGAGGCGGCCGCCAA 180
Qy 181 GGCCATCGGCGCGCTCAGGTGGAGCGCTATGTGCCCGAACAACAGGAGTCTGCTGGTG 240
Db 181 GGCCATCGGCGCGCTCAGGTGGAGCGCTATGTGCCCGAACAACAGGAGTCTGCTGGTG 240

RESULT 5

US-60-545-213-5825
; Sequence 5825, Application US/60545213

Qy 241 CCTGTGCTGCGGCTGTGAGGTGCGGGAAACACTGAGCCATGGAACCTGTGCGTGTCTGTA 300
Db 241 CCTGTGCTGCGGCTGTGAGGTGCGGGAAACACTGAGCCATGGAACCTGTGCGTGTCTGTA 300
Qy 301 CGGGGGGCTGCTGAGGATCTGGCCAGCCAGAGCAACAAGAACCAACAATTCCTG 360
Db 301 CGGGGGGCTGCTGAGGATCTGGCCAGCCAGAGCAACAAGAACCAACAATTCCTG 360
Qy 361 GTGGGAGAAACAAAGCTGAGGTCCAGATGAAAGAAAGTTTCTGTGCTCACTCCCCAGGATTA 420
Db 361 GTGGGAGAAACAAAGCTGAGGTCCAGATGAAAGAAAGTTTCTGTGCTCACTCCCCAGGATTA 420
Qy 421 TGCCCGATTCAAGAAATCCATGTGTGAAAGTTTGGATTTCCTATGAAGAAAGAGGATAA 480
Db 421 TGCCCGATTCAAGAAATCCATGTGTGAAAGTTTGGATTTCCTATGAAGAAAGAGGATAA 480
Qy 481 AGTGATCAAGGAGATGCGAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGGAGGTGGT 540
Db 481 AGTGATCAAGGAGATGCGAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGGAGGTGGT 540
Qy 541 TCGGCTCTGTCTTAGAGCTCAGGCACTGCGAGAGCCAGAGAGGGCTCTTCAGACACCTAG 600
Db 541 TCGGCTCTGTCTTAGAGCTCAGGCACTGCGAGAGCCAGAGAGGGCTCTTCAGACACCTAG 600
Qy 601 AAGCTGAAAGGGATGAACAGCCAAAGTAGTTCAGCTTACAGAGCCCTCAAAATTTGGA 660
Db 601 AAGCTGAAAGGGATGAACAGCCAAAGTAGTTCAGCTTACAGAGCCCTCAAAATTTGGA 660
Qy 661 CCTGCCACCACTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGG 720
Db 661 CCTGCCACCACTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGG 720
Qy 721 CCATCAGGATATACAGGAGTTGGTAACATCCACTCAGGTGCGACACCTCCTCGATGAT 780
Db 721 CCATCAGGATATACAGGAGTTGGTAACATCCACTCAGGTGCGACACCTCCTCGATGAT 780
Qy 781 CCAAGATGAAGATPACATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCT 840
Db 781 CCAAGATGAAGATPACATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCT 840
Qy 841 TAAAGAAAGGAAAGAAACAGAGTTGAAAGAACTCCCCCCAGACGAGTTGGGGCAACTT 900
Db 841 TAAAGAAAGGAAAGAAACAGAGTTGAAAGAACTCCCCCCAGACGAGTTGGGGCAACTT 900
Qy 901 TGATCACAGCTCCAGGACCAAGTGCAGGCTGCTGCTCTTTTGGGCGGCTCTGGAATA 960
Db 901 TGATCACAGCTCCAGGACCAAGTGCAGGCTGCTGCTCTTTTGGGCGGCTCTGGAATA 960
Qy 961 ATGGAGCGCGCTGGCAGTCCAGACATCAATCCAAAACCTGAAGTGCAGCAATGAAGAAG 1020
Db 961 ATGGAGCGCGCTGGCAGTCCAGACATCAATCCAAAACCTGAAGTGCAGCAATGAAGAAG 1020
Qy 1021 CAGTCACATACAGAAAAAGCTAATCATGTCTCTCAACAATCAATGAGGCTAAAAGCC 1080
Db 1021 CAGTCACATACAGAAAAAGCTAATCATGTCTCTCAACAATCAATGAGGCTAAAAGCC 1080
Qy 1081 AAAGTCAACCAACCCCTAATATACCTTCCACCCAAATCTTTATCATGTTCTTCTTAG 1140
Db 1081 AAAGTCAACCAACCCCTAATATACCTTCCACCCAAATCTTTATCATGTTCTTCTTAG 1140
Qy 1141 GAAACAGACATCACTCATTCATTGATTTAATAAAGTTTATTTTCGGCCTTCGTGGCCT 1200
Db 1141 GAAACAGACATCACTCATTCATTGATTTAATAAAGTTTATTTTCGGCCTTCGTGGCCT 1200
Qy 1201 CGAATCAAGCTTATCGATACCGTCCAGCTCGAGCTGGGGGGGGGCTACCCACTTTT 1253
Db 1201 CGAATCAAGCTTATCGATACCGTCCAGCTCGAGCTGGGGGGGGGCTACCCACTTTT 1253

```
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042099)
; CURRENT APPLICATION NUMBER: US/60/545,213
; CURRENT FILING DATE: 2004-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5825
; LENGTH: 1253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: n is a, c, g, or t
US-60-545-213-5825

Query Match      99.9%; Score 1252; DB 124; Length 1253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CATTGGAGTCACGGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60
DB
QY      1  CATTGGAGTCACGGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60
DB
QY      61  GATTTCAGGCCCAAGAGGCCCTCTCTGTGTGGTGGCGGCAGCTTTACAGCCGCAAGCAC 120
DB
QY      61  GATTTCAGGCCCAAGAGGCCCTCTCTGTGTGGTGGCGGCAGCTTTACAGCCGCAAGCAC 120
DB
QY      121  CCAGCGGCACTGAAGAGGCTTTTTCAGAGGCTCTCTGTGTGGTGGCGGCAGCTTTTACAGCCGCAAGCAC 180
DB
QY      121  CCAGCGGCACTGAAGAGGCTTTTTCAGAGGCTCTCTGTGTGGTGGCGGCAGCTTTTACAGCCGCAAGCAC 180
DB
QY      181  GGCATCCCGCGCTCAGGTGAGCGCTATGTGCCGGAACACGAGCGATGCTGTGGTG 240
DB
QY      181  GGCATCCCGCGCTCAGGTGAGCGCTATGTGCCGGAACACGAGCGATGCTGTGGTG 240
DB
QY      241  CCTGTGTGGTGGTGTGAGTGGCGGAACACCTTGAGCCATGGAACCTGACGGTGTGTA 300
DB
QY      241  CCTGTGTGGTGGTGTGAGTGGCGGAACACCTTGAGCCATGGAACCTGACGGTGTGTA 300
DB
QY      301  CGGGGGGCTCTCGAGCATCTGCCAGCCAGCAGCAAGAACCAACCAAAATTCG 360
DB
QY      301  CGGGGGGCTCTCGAGCATCTGCCAGCCAGCAGCAAGAACCAACCAAAATTCG 360
DB
QY      361  GTGGGAAACAAAGCTCAGGTCCAGATGAAAGAGAGTTTCTGTCTCACTCCCCAGGATTA 420
DB
QY      361  GTGGGAAACAAAGCTCAGGTCCAGATGAAAGAGAGTTTCTGTCTCACTCCCCAGGATTA 420
DB
QY      421  TGGCGGATTCAAGAAATCCATGTTGAAAGTTTGGATTCTTATGAAGAAAGAGGATTA 480
DB
QY      421  TGGCGGATTCAAGAAATCCATGTTGAAAGTTTGGATTCTTATGAAGAAAGAGGATTA 480
DB
QY      481  ACTGATCAAGAGATGTCAGCTCAGATCCAGATGAAAGAGAGTTTCTGTCTCACTCCCCAGGATTA 540
DB
QY      481  ACTGATCAAGAGATGTCAGCTCAGATCCAGATGAAAGAGAGTTTCTGTCTCACTCCCCAGGATTA 540
DB
QY      541  TCGGTCTGTCTTAGAGCTCAGGAGTCCAGAGTCCAGAGTCCAGAGAGGGCTCTTTCAGCACCTAG 600
DB
QY      541  TCGGTCTGTCTTAGAGCTCAGGAGTCCAGAGTCCAGAGTCCAGAGAGGGCTCTTTCAGCACCTAG 600
DB
QY      601  AAGCTGGAAGAGGATGAACAGCCAAAGTAGCTTCCAGCTTACAGAGCCCTCAAAATTTGGA 660
DB
QY      601  AAGCTGGAAGAGGATGAACAGCCAAAGTAGCTTCCAGCTTACAGAGCCCTCAAAATTTGGA 660
DB
QY      661  CCTGCCACCACTCCAGAGCTTGTACTGTGAGACAGAGCCATCTCTGACATTCATTGG 720
DB
QY      661  CCTGCCACCACTCCAGAGCTTGTACTGTGAGACAGAGCCATCTCTGACATTCATTGG 720
DB
QY      721  CCATCAGGATATACAGGAGTTGGTAAACATCCACTCAGGTGCCACACCTCCCTGGATGAT 780
```

```
DB      721  CCATCAGGATATACAGGAGTTGGTAAACATCCACTCAGGTGCCACACCTCCCTGGATGAT 780
QY      781  CCAAGATGAAGATACATTCCTGGGAAACCAAGAAATAGGACCATCTTATGAAGAAATTC 840
DB      781  CCAAGATGAAGATACATTCCTGGGAAACCAAGAAATAGGACCATCTTATGAAGAAATTC 840
QY      841  TAAAGAAAAGGAAAACAGAAAGTTGAAAAAACTCCCCCCAGAGTTCGGGCGCAACT 900
DB      841  TAAAGAAAAGGAAAACAGAAAGTTGAAAAAACTCCCCCCAGAGTTCGGGCGCAACT 900
QY      901  TGATCAGCTCCAGGACCAAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGCTGGAATA 960
DB      901  TGATCAGCTCCAGGACCAAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGCTGGAATA 960
QY      961  ATGGACCGCTGGCAGTCCAGACATCAACTCCAAAACCTGAAGCTGCAGCAATGAAGAAG 1020
DB      961  ATGGACCGCTGGCAGTCCAGACATCAACTCCAAAACCTGAAGCTGCAGCAATGAAGAAG 1020
QY      1021  CAGTCACATACAGAAAAAGCTAATCATGTCTCTACCAACTACCATGAGGCTAAAAGCC 1080
DB      1021  CAGTCACATACAGAAAAAGCTAATCATGTCTCTACCAACTACCATGAGGCTAAAAGCC 1080
QY      1081  AAGTCAACCAAAACCCCTATTATACCTTCCACCCAAATCTTTATCATTTCTTCTTAG 1140
DB      1081  AAGTCAACCAAAACCCCTATTATACCTTCCACCCAAATCTTTATCATTTCTTCTTAG 1140
QY      1141  GAAACAGACATACATTCATTTGATTAAAGTTTATTTTCGGGCTTCGTGGCCT 1200
DB      1141  GAAACAGACATACATTCATTTGATTAAAGTTTATTTTCGGGCTTCGTGGCCT 1200
QY      1201  CGAATCAAGCTTATCGATACCGTCGAGGGGGGGCGTACCCACTTTT 1253
DB      1201  CGAATCAAGCTTATCGATACCGTCGAGGGGGGGCGTACCCACTTTT 1253

RESULT 6
US-60-545-213-5826
; Sequence 5826, Application US/60545213
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042099)
; CURRENT APPLICATION NUMBER: US/60/545,213
; CURRENT FILING DATE: 2004-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5826
; LENGTH: 1253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: n is a, c, g, or t
US-60-545-213-5826

Query Match      99.9%; Score 1252; DB 124; Length 1253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CATTGGAGTCACGGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60
DB      1  CATTGGAGTCACGGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60
QY      61  GATTTCAGGCCCAAGAGGCCCTCTCTGTGTGGTGGCGGCAGCTTTACAGCCGCAAGCAC 120
DB      61  GATTTCAGGCCCAAGAGGCCCTCTCTGTGTGGTGGCGGCAGCTTTACAGCCGCAAGCAC 120
QY      121  CCAGCGGCACTGAAGAGGCTTTTTCAGAGGCTCTCTGTGTGGTGGCGGCAGCTTTTACAGCCGCAAGCAC 180
```


Db 121 CCAGGGCAGCTGAAGGAGGCTTTTGAAGAGGCTCTTGTCCCGCCAGGTGAGGGCGGCCGCA 180
Qy 181 GCCCATCCGCGCGCTCAGAGTGAAGCTATGTGCCCCGAAACACGAGGAGTCTGCTGGTG 240
Db 181 GCCCATCCGCGCGCTCAGAGTGAAGGCTATGTGCCCCGAAACACGAGGAGTCTGCTGGTG 240
Qy 241 CTTGTCTGCGGCTGTGAGGTGCGGGAAACACTGAGCCATGGAACCTGACGGTCTGTGA 300
Db 241 CTTGTCTGCGGCTGTGAGGTGCGGGAAACACTGAGCCATGGAACCTGACGGTCTGTGA 300
Qy 301 CGGGGGGCTGCTGAGGATCTGGCCAGCCAGAGCACAAGAAAGCAACCAACAATCTG 360
Db 301 CGGGGGGCTGCTGAGGATCTGGCCAGCCAGAGCACAAGAAAGCAACCAACAATCTG 360
Qy 361 GTGGGAGAACAAAGCTGAGGTCAGATGAAAGAAAGTTTCTGTGTCATCTCCCGAGGATTA 420
Db 361 GTGGGAGAACAAAGCTGAGGTCAGATGAAAGAAAGTTTCTGTGTCATCTCCCGAGGATTA 420
Qy 421 TGCGCGATTCAAGAAATCCATGGTGAAGTTTGGATTCTTATGAAGAAAGGAGGATAA 480
Db 421 TGCGCGATTCAAGAAATCCATGGTGAAGTTTGGATTCTTATGAAGAAAGGAGGATAA 480
Qy 481 AGTCAATCAAGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGGAGTGGT 540
Db 481 AGTCAATCAAGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGGAGTGGT 540
Qy 541 TCGGTCTGTCTTAGAGCTCAGGAGTGCAGACCCAGAGAGGGCTCTTCAGCACCTAG 600
Db 541 TCGGTCTGTCTTAGAGCTCAGGAGTGCAGACCCAGAGAGGGCTCTTCAGCACCTAG 600
Qy 601 AAGCTGGAAGGGATGAACAGCAAGTAGTTTCAGCTTACAGCAGCCCTCAAAATTTGGA 660
Db 601 AAGCTGGAAGGGATGAACAGCAAGTAGTTTCAGCTTACAGCAGCCCTCAAAATTTGGA 660
Qy 661 CCTGCCACCAGCTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATGG 720
Db 661 CCTGCCACCAGCTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATGG 720
Qy 721 CCATCAGATATACAGGAGTTGGTAAATCCATCCATCAGGTGCCACACCTCCCTGGATGAT 780
Db 721 CCATCAGATATACAGGAGTTGGTAAATCCATCCATCAGGTGCCACACCTCCCTGGATGAT 780
Qy 781 CCAAGATGAAGAATACATTTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAATTTCT 840
Db 781 CCAAGATGAAGAATACATTTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAATTTCT 840
Qy 841 TAAAGAAAGGAAAAACAGAGTTGAAAAAACTCCCGCCAGACCGAGTTGGGGCCAACTT 900
Db 841 TAAAGAAAGGAAAAACAGAGTTGAAAAAACTCCCGCCAGACCGAGTTGGGGCCAACTT 900
Qy 901 TGATCAGCTCCAGGACCAAGTGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGGAATA 960
Db 901 TGATCAGCTCCAGGACCAAGTGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGGAATA 960
Qy 961 ATGGACCGCGCTGGCAGTCCAGACATCAACTCCAAAACTGAAGCTGCAGCAATGAAGAAG 1020
Db 961 ATGGACCGCGCTGGCAGTCCAGACATCAACTCCAAAACTGAAGCTGCAGCAATGAAGAAG 1020
Qy 1021 CAGTCAACATACAGAAAAAAGCTATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCC 1080
Db 1021 CAGTCAACATACAGAAAAAAGCTATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCC 1080
Qy 1081 AAAGTCAACAAACCCCTATTATACCTTCCACCCAAATCTTTTATCATGCTTTCTTAG 1140
Db 1081 AAAGTCAACAAACCCCTATTATACCTTCCACCCAAATCTTTTATCATGCTTTCTTAG 1140
Qy 1141 GAAACAGACATACATTCATTGATTTAATAAAGTTTATTTTTCGGCCCTTCGTGGCCT 1200
Db 1141 GAAACAGACATACATTCATTGATTTAATAAAGTTTATTTTTCGGCCCTTCGTGGCCT 1200
Qy 1201 CGAATCAAGCTTATCGATACCGTGAACCTCGAGGGGGGGCGGTACCCACTTTT 1253
Db 1201 CGAATCAAGCTTATCGATACCGTGAACCTCGAGGGGGGGCGGTACCCACTTTT 1253

RESULT 7

US-09-471-275-1856/c
; Sequence 1856, Application US/09471275
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Contigs Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 782
; CURRENT APPLICATION NUMBER: US/09/471,275
; EARLIER FILING DATE: 1999-12-23
; EARLIER APPLICATION NUMBER: US 09/235,076
; EARLIER FILING DATE: 1999-01-20
; EARLIER APPLICATION NUMBER: US 09/234,611
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: US 09/240,371
; EARLIER FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: US 09/277,227
; EARLIER FILING DATE: 1999-03-25
; EARLIER APPLICATION NUMBER: US 09/271,490
; EARLIER FILING DATE: 1999-03-18
; EARLIER APPLICATION NUMBER: US 09/293,972
; EARLIER FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: US 09/274,861
; EARLIER FILING DATE: 1999-03-23
; EARLIER APPLICATION NUMBER: US 60/125,453
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: US 60/126,605
; EARLIER FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: US 09/306,350
; EARLIER FILING DATE: 1999-05-07
; EARLIER APPLICATION NUMBER: US 09/399,720
; EARLIER FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: US 09/404,284
; EARLIER FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 10451
; SOFTWARE: pt CT_genes Version 1.0
; SEQ ID NO 1856
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1201)
; OTHER INFORMATION: n = a,t,c or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1135)...(127)
; OTHER INFORMATION: similar to gi4455233 in the genept database release 114,
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters

US-09-471-275-1856

Query Match 81.5%; Score 1021.8; DB 22; Length 1200;
Best Local Similarity 98.7%; Pred. No. 4.5e-262;
Matches 1093; Conservative 0; Mismatches 7; Indels 7; Gaps 6;
Qy 81 CCTTTCTTCTGTGTGTCGCGC-GGCAGCTTTTACAGCCGCAAGCACCCAGCGCAGCTGAAGGAG 139
Db 1115 CCTTTCTTCTGTGTGCGGGCAGCTTTTACAGCCGCAAGCA-CCAGCGCAGCTGAAGGAG 1057
Qy 140 GCCTTTTACAGAGCTCTCTCCCGCAGGTGAGCGCCGCAAGGCATCCGCGCGCTCAG 199
Db 1056 GCCTTTTACAGAGCTCTCTCCCGCAGGTGAGCGCCGCAAGGCATCCGCGCGCTCAG 997
Qy 200 GTGAGCGCTATGTGCCCGCAACACGAGCGATGCTGTGCTGTGCTGTGCTGTGCTGTGAG 259
Db 996 GTGAGCGCTATGTGCCCGCAACACGAGCGATGCTGTGCTGTGCTGTGCTGTGCTGTGAG 937
Qy 260 GTGCGGGAACACCTTGAGCCCATGGAACCTCAACGCTGTGTACCGGGGGGCTGCTGGAGCAT 319
Db 1115 GTGCGGGAACACCTTGAGCCCATGGAACCTCAACGCTGTGTACCGGGGGGCTGCTGGAGCAT 319


```
Qy 500 GCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGAGAGGTGGTTCGGTCTGTCTTAGAGCCT 559
Db 696 GCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGAGAGGTGGTTCGGTCTGTCTTAGAGCCT 637
Qy 560 CAGCAGTGGCCAGACCCAGAGAGGGTCTTTACACCTAGAACCTGAGAGGATGAAC 619
Db 636 CAGCAGTGGCCAGACCCAGAGAGGGTCTTTACACCTAGAACCTGAGAGGATGAAC 577
Qy 620 AGCCAAGTAGCTTCCAGCTTACAGAGCCCTCAAATTTGGACCTGCCACAGCTCCAGAG 679
Db 576 AGCCAAGTAGCTTCCAGCTTACAGAGCCCTCAAATTTGGACCTGCCACAGCTCCAGAG 517
Qy 680 CTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGGCCCATCAGGATATACAGGA 739
Db 516 CTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGGCCCATCAGGATATACAGGA 457
Qy 740 GTTGGTAACATCCACTCAGGTGCCACCTCCCTGGATGATCCAAAGATGAAGATACATT 799
Db 456 GTTGGTAACATCCACTCAGGTGCCACCTCCCTGGATGATCCAAAGATGAAGATACATT 397
Qy 800 GCTGGGAACCAAGAAATAGGACCATCTCTATGAAGATTTCTTAAAGAAAGGAAACAG 859
Db 396 GCTGGGAACCAAGAAATAGGACCATCTCTATGAAGATTTCTTAAAGAAAGGAAACAG 337
Qy 860 AAGTTGAAAAAATCTCCCCCAAGACCGAGTTGGGCGCAACTTTTGATCAAGCTCCAGGACC 919
Db 336 AAGTTGAAAAAATCTCCCCCAAGACCGAGTTGGGCGCAACTTTTGATCAAGCTCCAGGACC 277
Qy 920 AGTCAGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGAAATATGAGCGCGCTGGCAGTC 979
Db 276 AGTCAGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGAAATATGAGCGCGCTGGCAGTC 218
Qy 980 CAGACATCAACTCCAAAATGAGCTGCAGCAATGAAGAAGCAGTCAATACAGAAAAA 1039
Db 217 CAGACATCAA--TTCAAACTGAGCTGCAGCAATGAAGAAGCAGTCAATACAGAAAAA 159
Qy 1040 GCTAATCATGCTCTTACCAACTACATGAGGCTAAAGCCAAAGTCAACCAACCCCTA 1099
Db 158 GCTAATCATGCTCTTACCAACTACATGAGGCTAAAGC--AAAGTCAACCAACCCCTA 101
Qy 1100 TTATACCTCCACCAAAATCTTTATCATGCTCTTTCTTAGGAAACAGACATCTCATTC 1159
Db 100 TTATACCTTCCA--CCAAATCTTTATCATGCTCTTTCTTAGGAAACAGACATCTCATTC 42
Qy 1160 ATTTGATTTAATAAGTTTATTTTC 1186
Db 41 ATTTGATTTAATAAGTTTATTTTC 15
```

RESULT 9

```
US-09-552-317-6991/c
; Sequence 6991, Application US/09552317
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Tillinghast, John
; APPLICANT: Sanku, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids
; TITLE OF INVENTION: and Polypeptides
; FILE REFERENCE: 784CIP
; CURRENT APPLICATION NUMBER: US/09/552,317
; CURRENT FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 10289
; SOFTWARE: pt CT_genes Version 1.01
; SEQ ID NO 6991
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (1)...(1201)
; OTHER INFORMATION: n = a,t,c or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1135)...(127)
; OTHER INFORMATION: similar to gi4455233 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-552-317-6991
```

```
Query Match      81.5%; Score 1021.8; DB 26; Length 1200;
Best Local Similarity 98.7%; Pred. No. 4.5e-262;
Matches 1093; Conservative 0; Mismatches 7; Indels 7; Gaps 6;
```

```
Qy 81 CTTCTTTCTGTGTGCG-GGCAGCTTTTACAGCCGCAAGCACCCACAGCGGAGCTGAAGGAG 139
Db 1115 CTTCTTTCTGTGTGCGGGCGACGTTTACAGCCGCAAGCA-CCAGCGGCACTGAAGGAG 1057
Qy 140 GCTTTTGAAGAGGCTCCTGCCCCCAGGTGGAGCGCGCCGCAAGCCATCCGCGCCGCTCAG 199
Db 1056 GCTTTTGAAGAGGCTCCTGCCCCCAGGTGGAGCGCGCCGCAAGCCATCCGCGCCGCTCAG 997
Qy 200 GTGGAGCGCTATGTGCCCGCAACACAGCGATGCTGTGTGTGCTGTGTGCTGTGTGCTGTG 259
Db 996 GTGGAGCGCTATGTGCCCGCAACACAGCGATGCTGTGTGTGCTGTGTGCTGTGTGCTGTG 937
Qy 260 GTGGGGAACACCTGAGGCCATGGAACCTGACGGTGTCTGTACGGGGGGCTGTGTGAGCAT 319
Db 936 GTGGGGAACACCTGAGGCCATGGAACCTGACGGTGTCTGTACGGGGGGCTGTGTGAGCAT 877
Qy 320 CTGGCCAGCCGACAGGACACAAAGAACCAACAAATTTCTGGTGGGAGAACAAAGCTCAG 379
Db 876 CTGGCCAGCCGACAGGACACAAAGAACCAACAAATTTCTGGTGGGAGAACAAAGCTCAG 817
Qy 380 GTCCAGATGAAGAGAGAGTTTCTGGTCACTCCCAGGATTTATGCGCATTCAGAAATCC 439
Db 816 GTCCAGATGAAGAGAGAGTTTCTGGTCACTCCCAGGATTTATGCGCATTCAGAAATCC 757
Qy 440 ATGTTGAAGGTTTGGATTCCTATGAAGAAAGAGGATTAAGTGATCAAGGAGATGSCA 499
Db 756 ATGTTGAAGGTTTGGATTCCTATGAAGAAAGAGGATTAAGTGATCAAGGAGATGSCA 697
Qy 500 GCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGAGAGTGGTTCGGTCTGTCTTAGAGCCT 559
Db 696 GCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGAGAGTGGTTCGGTCTGTCTTAGAGCCT 637
Qy 560 CAGCAGTGGCCAGACCCAGAGAGGGCTCTTTACACCTAGAACCTGAGAGGATGAAC 619
Db 636 CAGCAGTGGCCAGACCCAGAGAGGGCTCTTTACACCTAGAACCTGAGAGGATGAAC 577
Qy 620 AGCCAAGTAGCTTCCAGCTTACAGAGCCCTCAAATTTGGACCTGCCACAGCTCCAGAG 679
Db 576 AGCCAAGTAGCTTCCAGCTTACAGAGCCCTCAAATTTGGACCTGCCACAGCTCCAGAG 517
Qy 680 CTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGGCCCATCAGGATATACAGGA 739
Db 516 CTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGGCCCATCAGGATATACAGGA 457
Qy 740 GTTGGTAACATCCACTCAGGTGCCACCTCCCTGGATGATCCAAAGATGAAGATACATT 799
Db 456 GTTGGTAACATCCACTCAGGTGCCACCTCCCTGGATGATCCAAAGATGAAGATACATT 397
Qy 800 GCTGGGAACCAAGAAATAGGACCATCTCTATGAAGATTTCTTAAAGAAAGGAAACAG 859
Db 396 GCTGGGAACCAAGAAATAGGACCATCTCTATGAAGATTTCTTAAAGAAAGGAAACAG 337
Qy 860 AAGTTGAAAAAATCTCCCCCAAGACCGAGTTGGGCGCAACTTTTGATCAAGCTCCAGGACC 919
Db 336 AAGTTGAAAAAATCTCCCCCAAGACCGAGTTGGGCGCAACTTTTGATCAAGCTCCAGGACC 277
Qy 920 AGTCAGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGAAATATGAGCGCGCTGGCAGTC 979
Db 276 AGTCAGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGAAATATGAGCGCGCTGGCAGTC 218
```



```
QY 998 CTGAAGCTGAGCAATGAAGAGCAGTCAATACAGAAAAAGCTAATCATGCTCTCTAC 1057
Db 1223 CTGAAGCTGAGCAATGAAGAGCAGTCAATACAGAAAAAGCTAATCATGCTCTCTAC 1282
QY 1058 CAACCTACCTAGAGCTAAAGCCAAAGTCAACCAAAACCCCTATTATACCTTCCACCCAAA 1117
Db 1283 CAACCTACCTAGAGCTAAAGCC--AAAGTCAACCAAAACCCCTATTATACCTTCCA-CCAAA 1339
QY 1118 TTCTTTATCATGCTCTTTCTTAGGAACACAGACATCTCATTTGATTTAATAAGTT 1177
Db 1340 TTCTTTATCATGCTCTTTCTTAGGAACACAGACATCTCATTTGATTTAATAAGTT 1399
QY 1178 TTATTTTTC 1186
Db 1400 TTATTTTTC 1408

RESULT 13
US-60-213-362-1179/c
; Sequence 1179, Application US/60213362
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polymor
; TITLE OF INVENTION: Identified Thereby
; FILE REFERENCE: GX-0016 P
; CURRENT APPLICATION NUMBER: US/60/213,362
; CURRENT FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 8429
; SOFTWARE: PERL Program
; SEQ ID NO 1179
; LENGTH: 3215
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1000026.9
; NAME/KEY: unsure
; LOCATION: 1011-1042
; OTHER INFORMATION: a, t, c, g, or other
US-60-213-362-1179

Query Match 52.9%; Score 663.4; DB 90; Length 3215;
Best Local Similarity 95.1%; Pred. No. 7.2e-166;
Matches 750; Conservative 0; Mismatches 6; Indels 33; Gaps 5;

QY 426 GATTCAAGAAATCCATGGTGAAGGTTTGGATTCTCTATGAAGAAAGGAGGATAAAGTGA 485
Db 2587 GATTCAAGAAATCCATGGTGAAGGTTTGGATTCTCTATGAAGAAAGGAGGATAAAGTGA 2528
QY 486 TCAAGGAGATGCGAGCTCAGATCCGTGAGTGGAGCAGCGCAGAGGAGTGGTTCGGT 545
Db 2527 TCAAGGAGATGCGAGCTCAGATCCGTGAGTGGAGCAGCGCAGAGGAGTGGTTCGGT 2468
QY 546 CTGCTCTTAGAG-----CCTCAGGCGAGTGCAGACCCA 577
Db 2467 CTGCTCTTAGAGGTTGGTTTCCCTCGGAGGATCCAGACCACTCAGGCGAGTGCAGACCCA 2408
QY 578 GAAGAGGGCTCTTCAGCAGCTAGAGCTGGAAGGAGTGAACAGCAAGTAGCTTCCAGC 637
Db 2407 GAAGAGGGCTCTTCAGCAGCTAGAGCTGGAAGGAGTGAACAGCAAGTAGCTTCCAGC 2348
QY 638 TTACAGCAGCCCTCAAAATTGACCTGACCTGCCACAGCTCCAGAGCTTGACTGGATGAGACA 697
Db 2347 TTACAGCAGCCCTCAAAATTGACCTGACCTGCCACAGCTCCAGAGCTTGACTGGATGAGACA 2288
QY 698 GGACCATCTCTGACATTCATTGGCCATCAGGATATACAGGAGTGGTAAACATCCACTCA 757
Db 2287 GGACCATCTCTGACATTCATTGGCCATCAGGATATACAGGAGTGGTAAACATCCACTCA 2228
QY 758 GGTGCCACACCTCCCTGGATGATCCAAGATGAAGATAATTCCTGGGAAACCAAGAAATA 817
```

```
Db 2227 GTGGCCACACCTCCCTGGATGATCCAGATGAAGATACATTTGCTGGGACCCAGAAATA 2168
QY 818 GGACCATCTCTATGAAGAAATTTCTTAAAGAAAAAGAAAAACAGAGTTGAAAAAACTCCCC 877
Db 2167 GGACCATCTCTATGAAGAAATTTCTTAAAGAAAAAGAAAAACAGAGTTGAAAAAACTCCCC 2108
QY 878 CCAGACCGAGTTGGGGCCAACTTTTCATCACAGCTCCAGGACCCAGTGCAGGCTGGCTGCC 937
Db 2107 CCAGACCGAGTTGGGGCCAACTTTTCATCACAGCTCCAGGACCCAGTGCAGGCTGGCTGCC 2048
QY 938 TCTTTTGGGCCCGCTCTGGATAAATGGACCGCGCTGGCAGTCCAGACATCAACTCCAAA 997
Db 2047 TCTTTT-GGCCCGCTCTGGATAAATGGACCGCGCTGGCAGTCCAGACATCAA-TTCAAAA 1990
QY 998 CTGAAGCTGAGCAATGAAGAGCAGTCAATACAGAAAAAGCTAATCATGCTCTCTAC 1057
Db 1989 CTGAAGCTGAGCAATGAAGAGCAGTCAATACAGAAAAAGCTAATCATGCTCTCTAC 1930
QY 1058 CAACCTACCTAGAGCTAAAGCCAAAGTCAACCAAAACCCCTATTATACCTTCCACCCAAA 1117
Db 1929 CAACCTACCTAGAGCTAAAGCC--AAAGTCAACCAAAACCCCTATTATACCTTCCA-CCAAA 1873
QY 1118 TTCTTTATCATGCTCTTTCTTAGGAACACAGACATCTCATTTGATTTAATAAGTT 1177
Db 1872 TTCTTTATCATGCTCTTTCTTAGGAACACAGACATCTCATTTGATTTAATAAGTT 1813
QY 1178 TTATTTTTC 1186
Db 1812 TTATTTTTC 1804

RESULT 14
US-60-278-561-2059/c
; Sequence 2059, Application US/60278561
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; TITLE OF INVENTION: Polymorphisms Identified Thereby
; FILE REFERENCE: GX-0012-1 P
; CURRENT APPLICATION NUMBER: US/60/278,561
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 15598
; SOFTWARE: PERL Program
; SEQ ID NO 2059
; LENGTH: 3216
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1000026.9
; NAME/KEY: unsure
; LOCATION: 1011-1042
; OTHER INFORMATION: a, t, c, g, or other
US-60-278-561-2059

Query Match 52.9%; Score 663.4; DB 97; Length 3216;
Best Local Similarity 95.1%; Pred. No. 7.2e-166;
Matches 750; Conservative 0; Mismatches 6; Indels 33; Gaps 5;

QY 426 GATTCAAGAAATCCATGGTGAAGGTTTGGATTCTCTATGAAGAAAGGAGGATAAAGTGA 485
Db 2587 GATTCAAGAAATCCATGGTGAAGGTTTGGATTCTCTATGAAGAAAGGAGGATAAAGTGA 2528
QY 486 TCAAGGAGATGCGAGCTCAGATCCGTGAGTGGAGCAGCGCAGAGGAGTGGTTCGGT 545
Db 2527 TCAAGGAGATGCGAGCTCAGATCCGTGAGTGGAGCAGCGCAGAGGAGTGGTTCGGT 2468
QY 546 CTGCTCTTAGAG-----CCTCAGGCGAGTGCAGACCCA 577
```

```
Db 2467 CTGTCTTAGAGTTGGTTTCCCTCGGAGGATCCAGACCACTCAGGAGTGCACAGCCCA 2408
Qy 578 GAAGAGGCGCTCTTCAGCACCTAGAGCTGGAAGGGGATGAACAGCCCAAGTAGCTTCACGC 637
Db 2407 GAAGAGGCGCTCTTCAGCACCTAGAGCTGGAAGGGGATGAACAGCCCAAGTAGCTTCACGC 2348
Qy 638 TTACAGCAGCGCTCAAAATTTGGACCTGCCACCAGCTCCAGAGCTTGAAGTGGAGACA 697
Db 2347 TTACAGCAGCGCTCAAAATTTGGACCTGCCACCAGCTCCAGAGCTTGAAGTGGAGACA 2288
Qy 698 GGACCATCTCTGACATTCATTGGGCGCATCAGGATATACCAGAGTTGGTAAACATCACTCA 757
Db 2287 GGACCATCTCTGACATTCATTGGGCGCATCAGGATATACCAGAGTTGGTAAACATCACTCA 2228
Qy 758 GGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATATACATTGCTGGGAACCAAGAAATA 817
Db 2227 GGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATATACATTGCTGGGAACCAAGAAATA 2168
Qy 818 GGACCATCTCTGATGAAGATTTCTTAAGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAA 877
Db 2167 GGACCATCTCTGATGAAGATTTCTTAAGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAA 2108
Qy 878 CCAGACCGAGTTGGGCGCACTTTGATCAGAGTCCAGGACAGTGCAGGCTGCTGCC 937
Db 2107 CCAGACCGAGTTGGGCGCACTTTGATCAGAGTCCAGGACAGTGCAGGCTGCTGCC 2048
Qy 938 TCTTTTGGCGCGCTCTGGAATATGAGACGCGCTGGCAGTCCAGACATCAACTCCCAAAA 997
Db 2047 TCTTTT-GGCGCGCTCTGGAATATGAGACGCGCTGGCAGTCCAGACATCAA-TTCAAAA 1990
Qy 998 CTGAAGCTGACGCAATGAAGAGCAGTCAATACAGAAAAAGGCTAATCATGCTCTCTAC 1057
Db 1989 CTGAAGCTGACGCAATGAAGAGCAGTCAATACAGAAAAAGGCTAATCATGCTCTCTAC 1930
Qy 1058 CAATACCATGAGCTTAAGCCCAAGTCAACCAACCCCTATTATACCTTCCACCCAAA 1117
Db 1929 CAATACCATGAGGCTAAAGC--AAAGTCAACCAACCCCTATTATACCTTCCA-CCAAA 1873
Qy 1118 TTCTTTATCATGCTCTTTCTTAGGAAACAGACATCACTCATTCATTGATTTAATAAAGTT 1177
Db 1872 TTCTTTATCATGCTCTTTCTTAGGAAACAGACATCACTCATTCATTGATTTAATAAAGTT 1813
Qy 1178 TTATTTTC 1186
Db 1812 TTATTTTC 1804
```

```
RESULT 15
US-60-324-185-7629/c
; Sequence 7629, Application US/60324185
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF SEQUENCE POLYMORPHISMS USING
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCE DATABASES, AND SINGLE NUCLEOTIDE
; TITLE OF INVENTION: POLYMORPHISMS IDENTIFIED THEREBY
; FILE REFERENCE: GX-0019-1 P
; CURRENT APPLICATION NUMBER: US/60/324,185
; CURRENT FILING DATE: 2001-09-21
; NUMBER OF SEQ ID NOS: 35862
; SOFTWARE: PERL Program
; SEQ ID NO 7629
; LENGTH: 2561
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1301765.9
US-60-324-185-7629
```

Query Match 52.8%; Score 661.8; DB 102; Length 2561;
Best Local Similarity 94.9%; Pred. No. 1.7e-165;

```
Matches 749; Conservative 0; Mismatches 7; Indels 33; Gaps 5;
Qy 426 GATTCAAGAAATCCATGGTGAAGTTTGGATTCTTATGAAGAAAAAGAGGATAAAGTGA 485
Db 1933 GATTCAAGAAATCCATGGTGAAGTTTGGATTCTTATGAAGAAAAAGAGGATAAAGTGA 1874
Qy 486 TCAAGGAGATGGAGCTCAGATCCGTGAGTGGAGAGAGCCGACAGAGGTTGGTTCGGT 545
Db 1873 TCAAGGAGATGGAGCTCAGATCCGTGAGTGGAGAGAGCCGACAGAGGTTGGTTCGGT 1814
Qy 546 CTGTCTTAGAG-----CCTCAGGCGAGTGCACAGACCCA 577
Db 1813 CTGTCTTAGAGTTGGTTTCCCTCGGAGGATCCAGACCACTCAGGCGAGTGCACAGACCCA 1754
Qy 578 GAAGAGGCGCTCTTCAGCACCTAGAGCTGGAAGGGGATGAACAGCCCAAGTAGCTTCACGC 637
Db 1753 GAAGAGGCGCTCTTCAGCACCTAGAGCTGGAAGGGGATGAACAGCCCAAGTAGCTTCACGC 1694
Qy 638 TTACAGCAGCGCTCAAAATTTGGACCTGCCACCAGCTCCAGAGCTTGAAGTGGAGACA 697
Db 1693 TTACAGCAGCGCTCAAAATTTGGACCTGCCACCAGCTCCAGAGCTTGAAGTGGAGACA 1634
Qy 698 GGACCATCTCTGACATTCATTGGCCATCAGGATATACCAGAGTTGGTAAACATCACTCA 757
Db 1633 GGACCATCTCTGACATTCATTGGCCATCAGGATATACCAGAGTTGGTAAACATCACTCA 1574
Qy 758 GGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATATACATTGCTGGGAACCAAGAAATA 817
Db 1573 GGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATATACATTGCTGGGAACCAAGAAATA 1514
Qy 818 GGACCATCTCTGAGAAATTTCTTAAGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGG 877
Db 1513 GGACCATCTCTGAGAAATTTCTTAAGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGG 1454
Qy 878 CCAGACCGAGTTGGGCGCACTTTGATCAGAGTCCAGGACAGTGCAGGCTGCTGCC 937
Db 1453 CCAGACCGAGTTGGGCGCACTTTGATCAGAGTCCAGGACAGTGCAGGCTGCTGCC 1394
Qy 938 TCTTTTGGCGCGCTCTGGAATATGAGACGCGCTGGCAGTCCAGACATCAACTCCCAAAA 997
Db 1393 TCTTTT-GGCGCGCTCTGGAATATGAGACGCGCTGGCAGTCCAGACATCAA-TTCAAAA 1336
Qy 998 CTGAAGCTGACGCAATGAAGAGCAGTCAATACAGAAAAAGGCTAATCATGCTCTCTAC 1057
Db 1335 CTGAAGCTGACGCAATGAAGAGCAGTCAATACAGAAAAAGGCTAATCATGCTCTCTAC 1276
Qy 1058 CAATACCATGAGCTTAAGCCCAAGTCAACCAACCCCTATTATACCTTCCACCCAAA 1117
Db 1275 CAATACCATGAGGCTAAAGC--AAAGTCAACCAACCCCTATTATACCTTCCA-CCAAA 1219
Qy 1118 TTCTTTATCATGCTCTTTCTTAGGAAACAGACATCACTCATTCATTGATTTAATAAAGTT 1177
Db 1218 TTCTTTATCATGCTCTTTCTTAGGAAACAGACATCACTCATTCATTGATTTAATAAAGTT 1159
Qy 1178 TTATTTTC 1186
Db 1158 TTATTTTC 1150
```

Search completed: June 11, 2005, 04:01:07
Job time : 3143.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:40:00 ; Search time 63.5671 Seconds
(without alignments)
2537.148 Million cell updates/sec

Title: US-09-155-676B-5
Perfect score: 2190
Sequence: 1 IGVTWRRRSRIVDPRAAXNS.....RGLESSLISIPSTSGRGRTHF 417

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_16Dec04:*
- 1: Geneseqp1980s:*
 - 2: Geneseqp1990s:*
 - 3: Geneseqp2000s:*
 - 4: Geneseqp2001s:*
 - 5: Geneseqp2002s:*
 - 6: Geneseqp2003as:*
 - 7: Geneseqp2003bs:*
 - 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query Match		Length	DB	ID	Description
			%				
1	2178	99.5	417	2	AAW42401	AAW42401	TRAF2 bin
2	795	36.3	1115	5	AAB71230	AAB71230	Human leg
3	795	36.3	1115	7	ABW01535	ABW01535	Human lgs
4	795	36.3	1115	8	ADJ71905	ADJ71905	Human Lgs
5	670	30.6	156	4	ABG19580	ABG19580	Novel hum
6	554.5	25.3	741	4	ABG19581	ABG19581	Novel hum
7	503	23.0	312	4	ABG13680	ABG13680	Novel hum
8	309.5	14.1	65	4	AAM89643	AAM89643	Human imm
9	198	9.0	104	2	AAV16779	AAV16779	Human sec
10	196	8.9	104	7	ADA45163	ADA45163	Human pol
11	111.5	5.1	850	4	ABE68702	ABE68702	Drosophil
12	111	5.1	838	4	ABE64668	ABE64668	Drosophil
13	110.5	5.0	1878	8	ADI28627	ADI28627	Mycoplasma
14	109	5.0	4952	5	ADH47759	ADH47759	NOV18 pro
15	109	5.0	4952	7	ADP68294	ADP68294	Human NOV
16	109	5.0	4952	8	ADL25642	ADL25642	Human dia
17	109	5.0	5159	5	ADH48828	ADH48828	NOV47 pro
18	108.5	5.0	830	4	ABG08145	ABG08145	Novel hum
19	108.5	5.0	1225	3	AAW07973	AAW07973	A human c
20	108.5	5.0	1225	7	ADF09504	ADF09504	Human nat
21	108.5	5.0	1233	4	ABG04996	ABG04996	Novel hum
22	107.5	4.9	1516	5	ABP69840	ABP69840	Human pol
23	107	4.9	1419	5	ABP69842	ABP69842	Human pol
24	107	4.9	1879	8	ADI28615	ADI28615	Mycoplasma
25	106	4.8	270	4	AAB68357	AAB68357	Amino aci

26	105.5	4.8	850	8	ABM80887	Abm80887 Tumour-as
27	105.5	4.8	852	8	ABM80886	Abm80886 Tumour-as
28	105.5	4.8	1942	7	ADES9894	Ades9894 Human pro
29	105.5	4.8	2246	4	ABG05850	Abg05850 Novel hum
30	105	4.8	366	8	ADR86145	Adr86145 Aspergill
31	105	4.8	669	4	ABB69682	Abb69682 Drosophil
32	105	4.8	1770	6	ABO14791	Abol14791 Novel hum
33	105	4.8	2701	6	ABR92087	AbR92087 Human cer
34	105	4.8	2819	4	AAB35408	Aab35408 Human 07C
35	105	4.8	2819	8	ADQ97653	Adq97653 Human can
36	104.5	4.8	526	6	ADA54293	Ada54293 Human pro
37	104.5	4.8	526	7	ADG31710	Adg31710 Human pro
38	104.5	4.8	1094	5	ABP73717	Abp73717 Candida a
39	104	4.7	475	8	ADR41765	Adr41765 Protein s
40	104	4.7	475	8	ADS14358	Ads14358 Human 52-
41	103.5	4.7	578	2	AAW89273	Aaw89273 Granulocy
42	103.5	4.7	753	8	ADR14357	Adr14357 Human NF-
43	103.5	4.7	3111	4	ABB60327	Abb60327 Drosophil
44	103.5	4.7	3224	2	AAWS4235	Aaws4235 Human Nup
45	103.5	4.7	3224	6	ABO14790	Abol14790 Novel hum

ALIGNMENTS

RESULT 1
AAW42401
ID AAW42401 standard; protein; 417 AA.
XX
AC AAW42401;
XX
DT 15-APR-1998 (first entry)
XX
DE TRAF2 binding protein encoded by clone 15.
XX
KW Human tumour necrosis factor receptor-associated factor 2; TRAF2;
KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;
KW intracellular signalling activity; acute hepatitis;
KW autoimmune-induced cell death.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 18 /note= "not specified"
FT Misc-difference 320 /note= "not specified"
FT Misc-difference 338 /note= "not specified"
FT Misc-difference 356 /note= "not specified"
FT Misc-difference 358 /note= "not specified"
FT Misc-difference 388 /note= "not specified"
FT Misc-difference 388 /note= "not specified"
XX
PN WO9737016-A1.
XX
PD 09-OCT-1997.
XX
XX 01-APR-1997; 97WO-IL000117.
XX
XX 02-APR-1996; 96IL-00117800.
XX 26-AUG-1996; 96IL-00119133.
XX (YEDA) YEDA RES & DEV CO LTD.
XX Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;
XX WPI; 1997-503101/46.
XX N-FSDB; AAV03325.
XX
XX DNA encoding tumour necrosis factor receptor-associated factor binding

PT molecule - used for modulation or mediation in cells of the activity of
PT NF-KB.

XX Disclosure; Fig 5; 127pp; English.

XX The present sequence represents a TRAF2 binding protein, which is encoded
CC by clone 15. Clone 15 is a partial clone, which lacks most of its 5' end
CC of the coding DNA sequence. A cDNA library prepared from B-cells was
CC screened for proteins that associate with TRAF2, and clone 15 was
CC isolated. The clone 15 protein is capable of binding to at least amino
CC acids 222-501 of TRAF2. The TRAF-2 binding proteins can be used for
CC modulation or mediation in cells of the activity of NF-kappaB or any
CC other intracellular signalling activity modulated or mediated by TRAF2.
CC TRAF-binding proteins are especially used for prevention or treatment of
CC pathological conditions associated with NF-kB induction, e.g. acute
CC hepatitis, autoimmune-induced cell death, e.g. death of the beta
CC langerhans cells or the pancreas that results in diabetes, the death of
CC cells in graft rejection, the death of oligodendrocytes in the brain in
CC multiple sclerosis, and AIDS-inhibited T cell suicide which causes
CC proliferation of the AIDS virus and hence the AIDS disease. The proteins
CC are also useful for screening of ligands capable of binding to a protein,
CC which are useful for modulating cellular activity modulated/mediated by
CC TRAF2

XX SQ Sequence 417 AA;

Query Match 99.5%; Score 2178; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 6.2e-190;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IGVTWRSSRIVDPRAAXNSIRAHGPPFCGRGTTAASTQRLKEAFERLLPQVEAARK 60
Db 1 IGVTWRSSRIVDPRAAXNSIRAHGPPFCGRGTTAASTQRLKEAFERLLPQVEAARK 60
Qy 61 AIRAAQVERVYVPEHERCCWCLCCGCEVREHLSHGNTLVLYGGLLEHLSPEHKATNKPW 120
Db 61 AIRAAQVERVYVPEHERCCWCLCCGCEVREHLSHGNTLVLYGGLLEHLSPEHKATNKPW 120
Qy 121 WENKAEVOMKEKFLVTPQDYARFKGMVKGDSYEKEDKVIKEMAAQIREVEQSRQVY 180
Db 121 WENKAEVOMKEKFLVTPQDYARFKGMVKGDSYEKEDKVIKEMAAQIREVEQSRQVY 180
Qy 181 RSVLEPQAVDPPEGSSAPRSWKMNSQVASSLQQPSNLDLPAPELDMWETGPSLTFIG 240
Db 181 RSVLEPQAVDPPEGSSAPRSWKMNSQVASSLQQPSNLDLPAPELDMWETGPSLTFIG 240
Qy 241 HQDIPGVGNHSGATPPWMIQDEEYIAGNQEIGPSYEEFLKEKEKOKLKLPPDRVGANF 300
Db 241 HQDIPGVGNHSGATPPWMIQDEEYIAGNQEIGPSYEEFLKEKEKOKLKLPPDRVGANF 300
Qy 301 DHSRTSAGWLPSGPRLXWTPAVOTSTPKLQOXRSSHQKKAHLYOLPXGXP 360
Db 301 DHSRTSAGWLPSGPRLXWTPAVOTSTPKLQOXRSSHQKKAHLYOLPXGXP 360
Qy 361 KSTKPLLYLPKFFIIFLRKQTSYFIFXFNKVLFFGLRGLESSISIPSTSGRGRTHF 417
Db 361 KSTKPLLYLPKFFIIFLRKQTSYFIFXFNKVLFFGLRGLESSISIPSTSGRGRTHF 417

RESULT 2

AAB71230
ID AAB71230 standard; protein; 1115 AA.

XX AAB71230;
AC AAB71230;

XX 18-NOV-2002 (first entry)

DE Human legless homologue hlgs-1 partial protein.

XX Legless; human; lgs; Wnt/Wingless signaling pathway; Wnt; Wg;
KW tissue proliferation; tumour; cytoskeletal; cellular disorder; colon;
KW blood disorder; cancer; breast; head and neck cancer; brain; thyroid;
KW medulloblastoma; skin cancer; tissue regeneration; tissue repair.

XX OS

XX Homo sapiens.

XX US2002086986-A1.

XX PD 04-JUL-2002.

XX 27-JUL-2001; 2001US-00915543.

XX 28-JUL-2000; 2000US-0221502P.

XX (BASL/) BASLER K.

XX (BRUN/) BRUNNER E.

XX (PROE/) FROESCH B.

XX (KRAM/) KRAMPS T.

XX (PETE/) PETER O.

XX Basler K, Brunner E, Froesch B, Kramps T, Peter O;

XX WPI; 2002-635689/68.

XX N-PSDB; AAF88468.

Novel polypeptide useful in therapeutic method for treating disorders of cell fate such as cell differentiation or cell proliferation.

XX Example II; Fig 10B; 41pp; English.

XX This invention describes a novel polypeptide sharing one or more homologous amino acid domains with the legless (lgs) protein, a downstream component of the Wnt/Wingless (Wnt/Wg) signaling pathway involved in the formation and maintenance of spatial arrangements and proliferation of tissues during development, and in the formation and growth of many human tumours. The products of the invention have cytoskeletal activity and can be used to treat cellular disorders, blood disorders and cancers caused by over-stimulation of the Wnt pathway, where the cancerous condition is colon, breast, head and neck, brain, thyroid, medulloblastoma or skin cancer. The product could also be used to promote tissue regeneration and repair. This sequence represents the human legless (lgs) protein homologue hlgs-1 described in the disclosure of the invention

XX SQ Sequence 1115 AA;

Query Match 36.3%; Score 795; DB 5; Length 1115;

Best Local Similarity 55.6%; Pred. No. 7e-63; Indels 60; Gaps 8;

Matches 180; Conservative 17; Mismatches 67;

Qy 28 PFCGRGTTAASTQRLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87

Db 828 PFCGRGHVYSRKHQQLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 887

Qy 88 REHLSHGNTLVLYGGLLEHLSPEHKATNKPWENKAEVOMKEKFLVTPQDYARFKKSM 147

Db 888 REHLSHGNTLVLYGGLLEHLSPEHKATNKPWENKAEVOMKEKFLVTPQDYARFKKSM 947

Qy 148 VGLGDSYEKEDKVIKEMAAQIREVEQSRQVRSVLEPQAVPDPPEGSSAPRSWKMNS 207

Db 948 VGLGDSYEKEDKVIKEMAAQIREVEQSRQVRSVLE-----TGPR----- 990

Qy 208 QVASSLQOQSNLDLPAPELDMWETGPSLTFIGHQDIPGVGNHSGATPPWMIQDEEYIA 267

Db 991 -YALIVRSFAVLSR-----RTLKSGAFPPQ--TPEAHQ 1021

Qy 268 GNGEIGPSYEEFLKEKEKOKLKLPPDRVGANFDPHSSRTSAGWLPSFGPRLEXWTPAV 326

Db 1022 ARCLCAPRRGALKPPEPPTKLGVPPHTTRKARPHAATSP-----RPRCTRQAP--- 1072

Qy 327 QISTPKLKL---QOXRSSHIQKKA 347

Db 1073 -NKTQSLQLAGKARKTALHLQTKA 1095

RESULT 3

Db	1022	ARCLCAPRGALKPFPFGRTKLGVPPHPTTRKARPHAAKTS	1072
Qy	327	QTSTPKLKL--QQRSSHIQKKA	347
Db	1073	-NKTOSLQLAGKARKTALHLQTKA	1095
RESULT 4			
ADJ71905			
ID	ADJ71905	standard; protein; 1115 AA.	
XX	AC	ADJ71905;	
XX	AC		
XX	DT	20-MAY-2004 (first entry)	
XX	DE	Human Lgs/Bc19 partial polypeptide.	
XX	XX	Human; legless; lgs; cell differentiation disorder;	
XX	KW	cell proliferation disorder; cancer; Wnt pathway; medulloblastoma; colon;	
XX	KW	breast; head; neck; brain; thyroid; skin; blood disease;	
XX	KW	tissue regeneration; tissue repair; cytostatic; Lgs/Bc19.	
OS	OS	Homo sapiens.	
XX	XX	US2004038901-A1.	
XX	PN	26-FEB-2004.	
XX	PD		
XX	PF	22-SEP-2003; 2003US-00664859.	
XX	XX	28-JUL-2000; 2000US-0221502P.	
XX	PR	27-JUL-2001; 2001US-00915543.	
XX	PR		
XX	XX	(UYZU-) UNIV ZURICH.	
PA	PA		
XX	XX	Basler K, Brunner E, Froesch B, Kramps T, Peter O;	
PI	PI	WPI; 2004-203288/19.	
XX	DR	N-PSDB; ADJ71904.	
XX	DR		
XX	DR	Novel polypeptide sharing one or more homologue amino acid domains with	
Pt	Pt	Legless protein being functional homologue of Legless, useful for	
Pt	Pt	diagnosing disorders of cell fate.	
XX	XX		
PS	PS	Example 2; SEQ ID NO 17; 62pp; English.	
XX	XX		
CC	CC	The invention relates to a polypeptide sharing one or more homologous	
CC	CC	amino acid domains with a Legless (Lgs) protein and is therefore a	
CC	CC	functional homologue of Lgs. The invention also relates to a nucleotide	
CC	CC	sequence encoding a protein present in invertebrate and/or vertebrate	
CC	CC	organisms, the nucleotide sequence encoding a protein comprising a	
CC	CC	positive function in a regulatory pathway and the use of the polypeptide	
CC	CC	for the isolation of Lgs-binding proteins by carrying out an assay chosen	
CC	CC	from an in vitro binding assay with such a peptide or a co-	
CC	CC	immunoprecipitation from vertebrate or invertebrate cell lysates or a	
CC	CC	mammalian or yeast two hybrid assay. The polypeptide and polynucleotide	
CC	CC	are useful for treating disorders of cell fate, which involves	
CC	CC	administering therapeutic compounds chosen from invertebrate and	
CC	CC	vertebrate Lgs protein homologues or fragments, antibodies, antibody	
CC	CC	fragments, Lgs antisense DNA, Lgs antisense RNA, Lgs double-stranded RNA,	
CC	CC	small peptides or chemical and natural compounds being capable of	
CC	CC	interfering with Lgs function, synthesis and degradation. The disorders	
CC	CC	are related to cell differentiation or cell proliferation. The compound	
CC	CC	is administered to treat a cancerous condition by preventing progression	
CC	CC	from a pre-neoplastic or non-malignant condition to a neoplastic or	
CC	CC	malignant state. The cancerous condition is characterised by over-	
CC	CC	stimulation of the Wnt pathway and is medulloblastoma or cancer of the	
CC	CC	colon, breast, head and neck, brain, thyroid or skin. The therapeutic	
CC	CC	compound may also be administered to a blood disease to promote tissue	
CC	CC	regeneration and repair. This sequence represents a human Lgs/Bc19	
CC	CC	partial polypeptide of the invention.	
XX	XX		

```
SQ Sequence 1115 AA;
Query Match 36.3%; Score 795; DB 8; Length 1115;
Best Local Similarity 55.6%; Pred. No. 7e-63;
Matches 180; Conservative 17; Mismatches 67; Indels 60; Gaps 8;

QY 28 FFCGRGTTTAASTORQLKEAFERLLPQVEAARKAIRAAQVRYVPEHERCCWCLCCGCEV 87
DB 828 FFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVRYVPEHERCCWCLCCGCEV 887
QY 88 REHLHGHNLTVLYGGLLEHLASPEHKATNKFWENKAEVOMKEKFLVTPDYARFKKSM 147
DB 888 REHLHGHNLTVLYGGLLEHLASPEHKATNKFWENKAEVOMKEKFLVTPDYARFKKSM 947
QY 148 VGLGDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLFPQAVDPPEGSSAPRSWKGMNS 207
DB 948 VGLGDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLFPQAVDPPEGSSAPRSWKGMNS 990
QY 208 QVASSLQSPSNLDLPPAPPELDMWETGSLTFIGHQDIPGVGNIHSGATPPWMIQDEEYIA 267
DB 991 -VALTVRSPAVLSR-----RTLKSGAPPPQ--TPEAHPQ 1021
QY 268 GNGEIGSPSYEEFLKEKEKQK-LKLPDDRVRGANFDHSSRTSAGWLPSPGPRLEXWTPLAV 326
DB 1022 ARCLCAPRRGALKPEPPGRTLKLGVPDHTTRKARPHAAKTSP-----RPRCTROAP--- 1072
QY 327 QTSTPKLKL---QQRSSHIQKKA 347
DB 1073 -NKTQSLQLAGKARKTALHLQTKA 1095

RESULT 5
ABG19580
ID ABG19580 standard; protein; 156 AA.
XX AC ABG19580;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #19571.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS83767.
XX XX New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 49939; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
```

```
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 156 AA;
Query Match 30.6%; Score 670; DB 4; Length 156;
Best Local Similarity 77.9%; Pred. No. 1e-52;
Matches 134; Conservative 4; Mismatches 12; Indels 22; Gaps 4;

QY 6 WRRSRIVDPRAAXNSIRAHEGPFPCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAA 65
DB 5 WSRARL-QPOAPA-----AABGGF--GEAPAPGA-----EVEAARKAIRAA 42
QY 66 QVERYVPEHERCCWCLCCGCEVREHLHGHNLTVLYGGLLEHLASPEHKATNKFWENKA 125
DB 43 QVERYVPEHERCCWCLCCGCEVREHLHGHNLTVLYGGLLEHLASPEHKATNKFWENKA 102
QY 126 EVOMKEKFLVTPDYARFKKSMVKGLDSEYEKEDKVIKEMAAQIREVEQSRQ 177
DB 103 EVOMKEKFLVTPDYARFKKSMVKGLDSEYEKEDKVIKEMAAQIREVEQSRQ 154

RESULT 6
ABG19581
ID ABG19581 standard; protein; 741 AA.
XX AC ABG19581;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #19572.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS83768.
XX XX New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
```


DE Human immune/haematopoietic antigen SEQ ID NO:17236.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cystostatic; gene therapy; vaccine; metastasis.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
XX
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0203467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 01-SEP-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228927P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.


```
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX N-PSDB; AAK62424.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Claim 11; SEQ ID NO 17236; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention
XX
XX Sequence 65 AA;
SQ
Query Match 14.1%; Score 309.5; DB 4; Length 65;
Best Local Similarity 92.2%; Pred. No. 2.5e-20;
Matches 59; Conservative 0; Mismatches 2; Indels 3; Gaps 1;
Qy 238 FIGH---QDIPGVGNHSGATPPWMIQDEEVIAGNQSIGPSYEFLEKEKQKLLKLPD 294
Db 2 FFHFQLQDIPGVGNHSGATPPWMIQDEEVIAGNQSIGPSYEFLEKEKQKLLKLPD 61
Qy 295 RVGA 298
Db 62 RVGA 65
RESULT 9
AA16779
ID AA16779 standard; protein; 104 AA.
XX
XX AA16779;
XX
XX 27-JUL-1999 (first entry)
XX
XX Human secreted protein (clone as20_2).
XX
XX Secreted protein; human; tissue marker; genetic disease; gene therapy;
XX veterinary medicine; cell proliferation; immunostimulant; infection;
XX immunosuppressant; autoimmune disease; organ rejection; tumour; anaemia;
XX haematopoiesis; wound healing; fertility control; chemotaxis; analgesic;
XX thrombolytic; haemophilia; infarction; antimicrobial agent; cancer.
XX
XX Homo sapiens.
XX
XX WO9924469-A1.
XX
XX 20-MAY-1999.
XX
XX 06-NOV-1998; 98WO-US023829.
XX
```

```
PR 07-NOV-1997; 97US-00965789.
PR 04-NOV-1998; 98US-00185936.
XX
XX (GEMY ) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Evans C, Merberg D;
XX Treacy M, Agostino MJ;
XX
XX WPI; 1999-327362/27.
XX N-PSDB; AAX60579.
XX
XX Nucleic acid encoding secreted human proteins.
XX
XX Claim 8; Page 92; 107pp; English.
XX
XX The invention provides polynucleotides (AAX60579-X60687) encoding
XX specific secreted human proteins (AAY16779-Y16787). The nucleic acid
XX sequences are deposited under the accession number ATCC 98580. The
XX polynucleotides are used as tissue markers, chromosomal tags, for
XX diagnosis of genetic diseases, to generate anti-protein or anti-DNA
XX antibodies, also as nutritional sources and supplements and in gene
XX therapy. The secreted proteins are useful therapeutically, in human or
XX veterinary medicine, e.g. for modulating cell proliferation or
XX differentiation, as immunostimulants or immunosuppressants (for treating
XX infections, autoimmune disease, organ rejection, or to induce tumour
XX immunity), as regulators of haematopoiesis (e.g. for treating anaemia or
XX in conjunction with tumour therapy), to stimulate growth of tissue for
XX wound healing, as fertility control agents, for regulating chemotaxis or
XX chemokines (e.g. for directing cells to tumours or sites of infection), as
XX haemostatic and thrombolytic agents (e.g. in treatment of haemophilia or
XX infarctions), as antimicrobial agents, for modifying biorhythms,
XX appetite, or metabolism, as analgesics and many other uses. The proteins
XX are also used to raise antibodies, used as diagnostic immunoassay
XX reagents also (when neutralizing) for treating e.g. cancer
XX
XX Sequence 104 AA;
SQ
Query Match 9.0%; Score 198; DB 2; Length 104;
Best Local Similarity 62.5%; Pred. No. 7.9e-10;
Matches 40; Conservative 5; Mismatches 15; Indels 4; Gaps 1;
Qy 201 SWKGMNSQVASSLQPSNLDLPPAPELDWMETGSLTFIGHQDIPGVGNHSGATPPWMI 260
Db 19 AWMFFSQVASSLQPSNLDLPPAPELDWMETGSLTFIGHQ----VQRISKPEEGQWSL 74
Qy 261 QDEE 264
Db 75 RSQD 78
RESULT 10
ADA45163
ID ADA45163 standard; protein; 104 AA.
XX
XX ADA45163;
XX
XX 20-NOV-2003 (first entry)
XX
XX Human polypeptide #69.
XX
XX Human; genetic disorder; genetic fingerprinting; autoimmune disorder;
XX multiple sclerosis; systemic lupus erythematosus; graft-versus-host disease; anaemia;
XX insulin dependent diabetes mellitus; graft-versus-host disease; anaemia;
XX periodontal disease; bone fracture; cartilage damage;
XX central nervous system disorder; Alzheimer's disease;
XX Parkinson's disease; cancer; nutrition; carbon source; nitrogen source;
XX carbohydrate source.
XX
XX Homo sapiens.
XX
XX US2003044935-A1.
XX
XX 06-MAR-2003.
```

[illegible]

DB 630 EVNPRIPSSRDENAKFVLDKTEKNKYWQIYSSASPAFQNKWSLFGYYRYLLGLDPKQT 688

RESULT 14
ADH47759
ID ADH47759 standard; protein; 4952 AA.
XX
AC ADH47759;
XX
DT 25-MAR-2004 (first entry)
XX
XX NOV18 protein, SEQ ID 56.
DE
XX Antidiabetic; anorectic; cardiant; hypotensive; antiarteriosclerotic;
KW anorectic; virucide; antibacterial; fungicide; protozoacide; nootropic;
KW neuroprotective; antiParkinsonian; anticonvulsant; osteopathic;
KW antiarthritic; antiinflammatory; dermatological; antiasthmatic;
KW antilipemic; Gene therapy; human; metabolic disorder; diabetes; obesity;
KW viral infection; bacterial infection; fungal infection;
KW helminthic infection; protozoal infection; anorexia; cancer;
KW cardiovascular disease; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; epilepsy; immune disorder; haematopoietic disorder;
KW inflammatory skin disorder; asthma; dyslipidaemia; NOV18;
KW ARL-like protein; chromosome 12q12-q14.
XX
OS Homo sapiens.
XX
XX WO200268647-A2.
XX
XX 06-SEP-2002.
XX
XX 16-JAN-2002; 2002WO-US001311.
XX
XX 16-JAN-2001; 2001US-0261376P.
PR 18-JAN-2001; 2001US-0262454P.
PR 18-JAN-2001; 2001US-0262587P.
PR 31-JAN-2001; 2001US-0265530P.
PR 14-FEB-2001; 2001US-0268595P.
PR 28-FEB-2001; 2001US-0272409P.
PR 16-MAR-2001; 2001US-0276777P.
PR 17-MAY-2001; 2001US-0291672P.
PR 27-SEP-2001; 2001US-0325306P.
PR 18-OCT-2001; 2001US-0330336P.
PR 09-NOV-2001; 2001US-0345202P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Padigar M, Alsobrook JP, Colman SD, Spytek KA, Boldog F;
XX Vernet CAM, Li L, Shenoy S, Casman S, Guo X, Edinger S;
PI MacDougall J, Malyankar U, Patturajan M, Shinkets RA, Pena C;
PI Tchernev V, Zernhosen BD, Millett I, Miller C, Lepley DM, Smithson G;
PI Baumgartner J, Herrmann J, Peyman JA, Gorman L, Mezes P, Kekuda R;
PI Taupier RJ, Gerlach V, Grosse WM, Liu X, Ellerman K, Rothenberg M;
PI Stone DJ, Burgess CE;
XX
XX WPI; 2002-698671/75.
DR N-PSDB; ADH47758.
XX
XX New isolated NOVX polypeptides and polynucleotides, useful for
FT preventing, diagnosing or treating NOVX-associated disorders e.g.
FT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT asthma, or infections.
XX
XX Claim 1; Page 159-160; 380pp; English.
XX
XX The present invention relates to novel proteins (I) referred to as NOVX,
CC where x is any number from 1 to 18, and their coding sequences (II). The
CC proteins and their coding sequences are useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease,
CC preferably a NOVX-associated disorder such as metabolic disorders,
CC diabetes, obesity, infectious diseases (viral, bacterial, fungal,
CC helminthic, and protozoal), anorexia, cancer, cardiovascular diseases
CC (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's

CC disease, Parkinson's disease, epilepsy, immune disorders
CC (osteoarthritis), haematopoietic disorders, inflammatory skin disorders,
CC asthma, and various dyslipidaemias. NOV18 comprises a ARL-like protein
CC and maps to chromosome 12q12-q14.
XX
SQ Sequence 4952 AA;
Query Match 5.0%; Score 109; DB 5; Length 4952;
Best Local Similarity 20.9%; Pred. No. 32;
Matches 80; Conservative 51; Mismatches 149; Indels 102; Gaps 15;
Qy 101 GGLLEHLASPE-----HKATNKFWEKA-----EVQMKKFLVTPQDYAR 142
Db 2609 GGPPAHLTSPLSLPGSGSSLEKFELESGALTGCGPAASGDELDKMESSLVASE---- 2664
Qy 143 FKSMVKGLDSYEKKDKVKEMAAQIREVEQSRQEVRSVLPE-----QAVDPDEGSS 197
Db 2665 -LPLLIEDLLEHKKELQKKQLSAQLQPAQQQQQQQSHLLPAPGPAQAMSLPHEGSS 2723
Qy 198 APRSWKGMNSQVA-----SSLQSPSNLDLPPAPELDMMETGSPSLTFIGHQ----- 242
Db 2724 P--SLAGSQQLSLGLAVARQPGLPQPLMPTQPPAHALQ--QLAPSNMAMVSNQGHMLSGQ 2780
Qy 243 -----DIPGVGNTHSGATPPWMIQDEYIAGNQEIGPSY-----EEF 279
Db 2781 HGGQAGLVPOQSSQPVLSPKPMGTMPFSCMKPQQLAMQQLANSFPFDDLDKFAEDI 2840
Qy 280 LKEKEKOK-----LKKLPDRVGANPDHSSRTSAGWLPSPGPRLEXWTFPLAVQTSTPKL 333
Db 2841 IGPIAKAKMVALKGIKV-----MAQSGIVAPGNRQ-----QVSLLAQ 2880
Qy 334 KLOQXRSSTHQKKANHALYQLPXGKPKSTKPLLYLPKPF--IIVFLRKQTSFIFXFN- 390
Db 2881 RLSGSPSSDLQNHVAAGSGQERSAGDPSQRP---NPPTFAQGVINEADQORQVEEWLFHT 2937
Qy 391 -KVLFFGLRGLESLSIPSTSR 411
Db 2938 QQLQOMQLKVLLEQIGVHRKSR 2959
RESULT 15
ADP68294
ID ADP68294 standard; protein; 4952 AA.
XX
AC ADP68294;
XX
DT 12-AUG-2004 (first entry)
XX
XX Human NOV18a protein, an ARL like protein SeqID 56.
XX
XX human; NOVX; Alzheimer's disease; Huntington's; inflammatory;
KW Crohn's disease; rheumatoid arthritis; immunological; endocrine;
KW pigmentation; haematopoietic; psychotic; autoimmune; muscular;
KW osteoporosis; angina pectoris; hypotension; anxiety; bulimia;
KW cancer; manic depression; virucide; antibacterial; analgesic;
KW neuroprotective; nootropic; cerebroprotective; anticonvulsant;
KW dermatological; osteopathic; antiarthritic; antiinflammatory; cytostatic;
KW hypotensive; cardiant; hypertensive; antiulcer; antiallergic;
KW antianginal; immunosuppressive; antidepressant; neurodegenerative; ARL.
XX
XX Homo sapiens.
OS
XX WO200281510-A2.
XX
XX 17-OCT-2002.
XX
XX 18-JAN-2002; 2002WO-US001467.
XX
XX 18-JAN-2001; 2001US-0262454P.
PR 23-JAN-2001; 2001US-0263605P.
PR 25-JAN-2001; 2001US-0264159P.
PR 31-JAN-2001; 2001US-0265517P.
PR 07-FEB-2001; 2001US-0267057P.
PR

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:43:10 ; Search time 14.4085 Seconds
(without alignments)
2784.626 Million cell updates/sec

Title: US-09-155-676B-5
Perfect score: 2190
Sequence: 1 IGTVRRSRIVDPRAAXNS.....RGLESSUSIPSTSGRGTRHF 417

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000.

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	134	6.1	423	2 T05511	hypothetical prote
2	118.5	5.4	1479	2 T17401	transcription regu
3	110.5	5.0	673	2 S25335	transcription fact
4	109	5.0	4957	2 T03455	ALR protein - huma
5	109	5.0	5262	2 T03454	ALR protein - huma
6	106.5	4.9	850	2 T13352	stn-A protein - fr
7	105	4.8	544	2 AH2494	hypothetical prote
8	103.5	4.7	753	2 S48059	metal-regulatory t
9	103.5	4.7	1980	2 S54307	myosin heavy chain
10	103.5	4.7	3224	1 S56884	Ran-binding protei
11	103.5	4.7	3498	2 T22330	hypothetical prote
12	103	4.7	3449	2 T01083	hypothetical prote
13	102.5	4.7	1009	2 S60248	protein-tyrosine k
14	102	4.7	762	2 B86373	protein T23823.10
15	101	4.6	359	2 T01275	hypothetical prote
16	101	4.6	467	1 A49377	involucrin - mouse
17	101	4.6	560	2 S54093	hypothetical prote
18	101	4.6	2101	2 A42184	nuclear mitotic ap
19	100.5	4.6	771	1 A33430	h-caldesmon - chic
20	100.5	4.6	1277	2 S53043	probable membrane
21	100	4.6	810	2 G01252	small GTP binding
22	100	4.6	1027	2 T46481	hypothetical prote
23	100	4.6	1031	2 D88912	protein T06A10.1
24	100	4.6	1031	2 T33655	hypothetical prote
25	100	4.6	1106	2 T31742	hypothetical prote
26	99	4.5	475	1 A37241	52k autoantigen Ro
27	99	4.5	1805	2 A34736	nestin - rat
28	98.5	4.5	357	2 T17027	MYB-related transc
29	98.5	4.5	1009	2 A57434	protein-tyrosine k

RESULT 1

T05511
hypothetical protein F13M23.40 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: T05511
R/Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.
submitted to the Protein Sequence Database, February 1999
A/Reference number: Z15419
A/Accession: T05511
A/Molecule type: DNA
A/Residues: 1-423 <BEV>
A/Cross-references: UNIPROT:Q9SW36; EMBL:AL035523
A/Experimental source: cultivar Columbia; BAC clone F13M23
C/Genetics:
A/Map position: 4
A/Introns: 88/2; 130/3; 228/1; 253/3; 272/1
A/Note: F13M23.40
C/Superfamily: Arabidopsis thaliana hypothetical protein F13M23.40

ALIGNMENTS

Query Match		6.1%;	Score 134;	DB 2;	Length 423;
Best Local Similarity		18.2%;	Pred. No. 0.031;		
Matches		73;	Conservative	58;	Mismatches 153; Indels 116; Gaps 13;
Qy	44	LKEAFERLLPQVEAARKAIRAAOVRVVPHERCCWLCGCEVREHLHGNTLVLYGGL	103		
Db	36	LSSLLDFRSKIADVRFLLKNPSVLRPQSQNRVMVCFDEDIVELGS	91		
Qy	104	LEHLASPEHKKATNKFWENKAQVQMKKFLVTPQDYA	141		
Db	92	INHFASSDHLKNIKQFLSKNGKGPAMDCTDFRISADVAKWEKKQSGNEDASFEQSGQ	151		
Qy	142	-----RPFKSMVKGLDSYEEKEDKVIKEMAAQIREVQSGRQVRVSV	183		
Db	152	LSGTSNDIHTKLAFTMDRIKKVPAHHYSKNSD---VMPQYNTNTEYQISLSEIPGVI	208		
Qy	184	-----LEPQAVPDPEEGSSA-----PRSWKG-----MNSQVASSLQOQPSNLD	220		
Db	209	HNGSYLNMDDSQPLCDSENGSGFGEHSIPCRSKDSYNGNYCTOENTQVSDKKQIDGSY	268		
Qy	221	LPPAPELDMMETGSLTFIIGHQDIPGVGNHSGATPPWM-IQDEEYIAGNQEIGPSYBEF	279		
Db	269	NPPG-----VVGWTSISSSHSTDAG-GNVHSGAPPWLDANDGDF--SSVQLNQSDVAR	319		
Qy	280	LKSEKQKLLKLPDDRVGANFDH-----SSRTSAGWLPSFGPRLEXWT--	322		
Db	320	FOAKVPCKNRKLNPNRVGAWAERRKIEIEMKSGHVTKSNIDPDPLNPFGRVQSGTRK	379		
Qy	323	-----PLAVOTSTPKLQQRSSHIQKA	347		
Db	380	ESRKEFEKEKRLVKTESISTESEPVKIQ-----PYISKRA	415		

plasminogen-bindin
nestin - golden ha
hypothetical prote
sensory box sensor
hypothetical prote
calcium-binding pr
hypothetical prote
hypothetical prote
gravin - human
MADS box protein -
functional earcopl
DNA-binding protei
SNF2alpha protein
hypothetical prote
hypothetical prote
hypothetical prote

RESULT 2
T17401
transcription regulator WBSCR9 - mouse
N;Alternate names: Williams-Beuren syndrome deletion transcript 9 homolog
C;Species: Mus musculus (house mouse)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17401
R;Peoples, R.J.; Cisco, M.J.; Kaplan, P.; Francke, U.
Cytogenet. Cell Genet. 82, 238-246, 1998
A;Title: Identification of the WBSCR9 gene, encoding a novel transcriptional regulator,
A;Reference number: Z18735, MUID:9907764; PMID:9858827
A;Accession: T17401
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1479 <PEO>
A;Cross-references: UNIPROT:Q92277; EMBL:AF084480; NID:G4165088; PIDN:AA000000
C;Genetics:
A;Gene: Wbscr9
A;Map position: 5
F;1360-1415/Domain: bromodomain homology <BRO>

Query Match	5.4%;	Score 118.5;	DB 2;	Length 1479;
Best Local Similarity	21.1%;	Pred. No. 2.2;		
Matches	95;	Conservative	49;	Mismatches 118;
				Indels 189;
				Gaps 24;
Qy	35	FTAASTORQLKEAFERLLPQVEAARKAIRAAQAVERYVPEHERCCWCLCCGCEVREHL	SHG	94
Db	23	FTIPHTQ-----EAF-RTRERYE-----ARLRY-----SERIWTCKSTG-----	SS	58
Qy	95	NLTVLVGGILLEHLASPHSHKATNKFWENKAEEVQ--MKEKFLVTPQDYARFKSMV---	148	
Db	59	QLT-----HKEA-----WEEQEQVAELLKEEF--PNWYKLVLEWVHHT	96	
Qy	149	-----KGLDS-----YBEKEDVKIKEMAAQITREVEQSRQEVSVLEPQA	188	
Db	97	ASLEKLVDSAWLEIMTKYAVGEECDFEVGEKMLVKVIKVIHPLEK-----VDEEA	147	
Qy	189	VDPDEEGS-SAPRSWKGMNSQVASSLQQ-----PSNLD--	220	
Db	148	VEKSDGACDSPSDKENSQMAQDLQKETVVKVEDGRRESINDRARRSPRKLPTSLKK	207	
Qy	221	-----LP-----PAPELDWMETGPSLT-----FTGHQDIPGVGN	249	
Db	208	GERKWAPEKFLPHKYDVKLQNEDKIISNVFADSLIRTERPNKEILYFIHNNAL-----	R	263
Qy	250	IHSGATPPWMIQDEEYIAGNQIEGPSYEEFLKEKEK-----OKLKCLPPDPRVGANFD	301	
Db	264	AGTGENAPWVEDE--LVKKYSLPSKSDFLDDPKYKWTLPSTKRRNTGSPDRKPSK--	319	
Qy	302	HSRSTASGWLPSFGPRLEXTWPTLAVQSTSTPKLQLQXRSS-----	341	
Db	320	KPKRDSLSLSPNLNPKLVCHVHLEKSLNGPPLKYVKNKSNKSKSPEEHLEGVYKIMSPNNK	379	
Qy	342	-----HIQKXANHALYQLPXGKKP--KSTKPL	366	
Db	380	LHSPHIPKPG-----PAAKKPGKHSKDPL	403	

RESULT 3
S35335
transcription factor MTF-1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C;Accession: S35335; S36946
R;radtke, F.; Heuchel, R.; Georgiev, O.; Hergersberg, M.; Gariglio, M.; Dembic, Z.; Schaub
EMBO J. 12, 1355-1362, 1993
A;Title: Cloned transcription factor MTF-1 activates the mouse metallothionein I promoter
A;Reference number: S35335; MUID:93223676; PMID:8467794
A;Accession: S35335
A;Molecule type: mRNA
A;Residues: 1-673 <RAD>
A;Cross-references: UNIPROT:O9JJW8; EMBL:X71327

R;Radtke, F.
submitted to the EMBL Data Library, May 1993
A;Reference number: S36946
A;Accession: S36946
A;Molecule type: mRNA
A;Residues: 1-61,63-93,'Q',94-283,'N',285-673 <RA2>
A;Cross-references: EMBL:X71327
C;Genetics:
A;Gene: MTF-1

Query Match	5.08;	Score 110.5;	DB 2;	Length 673;
Best Local Similarity	21.3%;	Pred. No. 3;		
Matches	73;	Conservative	47;	Mismatches 134; Indels 89; Gaps 15;
Qy	21	IRAHEG--PPFC--	--GRGTTTAASTQQLKEAFERLLPQVEAAR	KAIRAAQVARYVPEH- 74
Db	249	IRTHTGKPPRCDHGGGKAFAAS		
Qy	75	-ERCCWCLCGCE-	-----VREHL--SHGNLTVLYGGLLEHLASPEHK-	-----114
Db	284	GERPFCPSNGCEKTFSTQVSLXSHMKGHDKNTAYSALPQHGSEDTNHSLYLSELGLL	343	
Qy	115	ATNKFWMENKAQVOMKEKFLVTPQDYARFKKSMVKGLDSEYEKEDKVIKEMAAQIREVEQ	174	
Db	344	STDSELQSSNSTQDQLSTISP--AIIFESM-----	PQNSDDPGIQDDPLQTAALID 394	
Qy	175	SRQEVRSVLEPOAVDPDPEEGSSAPRSWKGNMSQVASSLQPSNLIDLPPAPELDMWETGP	234	
Db	395	SFNGDAESVID--	-VPPP-----ACNSASLSLPLVLQSGISEPFPOLLP--ATAP 439	
Qy	235	SLTFIGHQDIPGVGNITHSGA-	-----TPPMWIMQDEEY-IAGNOBIGPSEYEFLEKESEKQK 287	
Db	440	SAP-----	-PPAPSLGPGSQAPAFGSPFALLQPPPEVPVPHSTQFAANHQEFLPHQAPP 492	
Qy	288	LKKLPDRVRGANDFHDSSRTSAGWLPSPGFPLEXTWTPLAQVTS	330	
Db	493	OTIVPGLSVVAGAPASGAATVASAVAAPPOSTTEPLPAMVOT	535	

RESULT 4
T03455
ALR protein - human
C/Species: Homo sapiens (man)
C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C/Accession: T03455
C/Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano, H.
Oncogene 15, 549-560, 1997
A/Title: Structure and expression pattern of human ALR, a novel gene with strong homology
A/Reference number: Z14954; MUID:97388474; PMID:9247308
A/Accession: T03455
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-4957 <PRA>
A/Cross-references: UNIPROT:O14686; EMBL:AF010404; NID:G2358286; PIDN:AAC51735.1; PID:G2358286
C/Genetics:
A/Name: ALR
A/Map position: 12
C/Superfamily: acute lymphoblastic leukemia protein, ALR type
C/Keywords: alternative splicing

	Query Match	5.0%;	Score 109;	DB 2;	Length 4957;
	Best Local Similarity	20.9%;	Pred.No. 55;		
	Matches	80;	Conservative	51;	Mismatches 149;
					Indels 102;
					Gaps 15;
Qy	101	GGLEHHLASPE-----HKATNKFWENKA-----	-----EVOMKEKFLVTPQDYAR	142	
Db	2609	GGPFAHLLTSPSGPGSSLLKFELESALTLPGGPAASGDDLDXWESSLVASE-----	26646		
Qy	143	FKKSMVKGLDYSBEKEDKVIKEMAAQIREVEQSRQEVVRSVLEB-----QAVPPPEEGSS	197		
Db	2665	-LPULLIDLEHBEKKELOKKQKQSAQLOPACQOQOQOQOQOQSHLLPAPGPAQMSLPHEGSS	2723		
Qy	198	APRSWKGMNSQVA-----SSLOQPSNLDLPPAPELDMWETGPSLTFTHGQ-----	242		

[illegible]

RESULT 6	Query 1	Base L	Match
T13352	stn-A pr		
C:Specie	A;Access		
C:Date:	A;Status		
C:Access	A;Molecu		
R;Kelly,	A;Residu		
submitte	A;Cross-		
A;Refer	C;Geneti		
A;Access	A;Cross-		
A;Status			
A;Molecu	Query 2	Base L	Match
A;Residu	stn-A pr		
A;Cross-	C:Specie		
C;Geneti	C:Date:		
A;Cross-	C:Access		
	R;Kelly,		
	submitte		
	A;Refer		
	A;Access		
	A;Status		
	A;Molecu		
	A;Residu		
	A;Cross-		
	C;Geneti		
	A;Cross-		
	Query 3	Base L	Match
	stn-A pr		
	C:Specie		
	C:Date:		
	C:Access		
	R;Kelly,		
	submitte		
	A;Refer		
	A;Access		
	A;Status		
	A;Molecu		
	A;Residu		
	A;Cross-		
	C;Geneti		
	A;Cross-		

7	RESULT
7	AH2494
7	hypothet
7	C;Specie
7	A;Note:
7	C;Date:
7	C;Access
7	R;kaneko
7	Nakazak
7	DNA Res.
7	A;Title:
7	A;Referer
7	A;Access
7	A;Status
7	A;Molecu
7	A;Residu
7	A;Cross-
7	A;Experi
7	C;Geneti
7	A;Gene:
7	A;Genome

A;Reference number: S58884; MUID:95327194; PMID:7603572
A;Accession: S58884
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-3224 <YOK>
A;Cross-references: UNIPROT:P49792; EMBL:D42063; NID:G924266; PIDN:BAA07662.1; PID:g1009
A;Experimental source: cell type B-lymphocyte
R;Wu, J.; Matunis, M.J.; Kraemer, D.; Blobel, G.; Coutavas, E.
J. Biol. Chem. 270, 14209-14213, 1995
A;Title: Nup358, a cytoplasmically exposed nucleoporin with peptide repeats, Ran-GTP bind
A;Reference number: A57545; MUID:95294031; PMID:7775481
A;Accession: A57545
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-776; R', 778-783, R', 785-3224 <WUA>
A;Cross-references: GB:L41840; NID:G857367; PIDN:AAC41758.1; PID:g857368
A;Experimental source: cell line HeLa
C;Genetics:
A;Gene: GDB:RANBP2; NUP358
A;Cross-references: GDB:4642758; OMIM:601181
A;Map position: 2cen-2ql3
C;Function:
A;Description: may play a role in nuclear protein import
C;Superfamily: nucleoporin Nup358; cyclophilin homology; tetratricopeptide repeat homolog
C;Keywords: leucine zipper
F;26-59/Domain: tetratricopeptide repeat homology <TT1>
F;60-93/Domain: tetratricopeptide repeat homology <TT2>
F;450-471/Domain: leucine zipper #strapped predicted <LEU>
F;3063-3224/Domain: cyclophilin homology <CYP>

Query Match 4.7%; Score 103.5; DB 1; Length 3224;
Best Local Similarity 18.2%; Pred. No. 79;
Matches 74; Conservative 69; Mismatches 132; Indels 131; Gaps 17;

Qy 48 PERLLPQVEAARKA-----IRAAQVERVVPHERCCWCLCGCEVRE 89
Db 606 WKKVLPLLIKKKNSIPEDIDPLFKHFSVDIOASEIVEY-----BEDA 650

Qy 90 HLSHGNLTVLYGGLLEHLASPEKKATNKFW-----WENKA-----EVQMKKF 133
Db 651 HITAILDAVNGNIEDAVTAFESIKSVSVYWNALIFHRKAEDIENDALSPEEQECKNY 710

Qy 134 LVTPQDVARFKSMVKGLDSVEEKEDKVIKMAAQIREVEOSROEVVRVLEPOAVPDPE 193
Db 711 LRKTRDY-----LIKIIDD-SDSNLSVVKLPVPLESV-----KEMLSNV- -QELSDYS 757

Qy 194 EGSSAPR--SWKGMNSQVASSLQOPSNDLPPAPELDWMTGPSLTFTGHQDIPGVGNH 251
Db 758 EGGPLYKNGSLRNADSEIKHSTPSTKYSJSPSKSYK-----Y 795

Qy 252 SGATPPPMWIDEEYIAGNQEIGPSYEFLEKEKOKL-----KKLPPDRVGA 298
Db 796 SPKTPPRAEDQNSLL--KMICQVEALKEMQBLKLNSSNSASPHRWPTENYGPSVD 853

Qy 299 NFDHSSRTSAGWLPSPGFRLEXWTPLAVQTSPTKLKIQXRSSHIOKKANHALYQLPXGX 358
Db 854 GY-QGSQTFHG-----APLTVATTGPSV-----VYSQSPAYNSQYLLRPA 893

Qy 359 KPKSTKLLY-----LPPKFFIIVFLRK-----QTYFIYFNKULF 394
Db 894 NVPTTKGPVYGMNRLPPQOHYIAYPQOMHTPPVQSSACMFQSBMY 939

RESULT 11
T22330
hypothetical protein F47A4.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22330
R;Mortimore, B.
submitted to the EMBL Data Library, June 1995
A;Reference number: Z19549
A;Accession: T22330

A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-3498 <MIL>
A;Cross-references: UNIPROT:Q20497; EMBL:Z49888; PIDN:CAA90064.1; GSPDB:GN00028; CESP:F4
A;Experimental source: clone F47A4
C;Genetics:
A;Gene: CESP:F47A4.2
A;Map position: X
A;Introns: 36/3; 72/3; 147/2; 228/3; 261/2; 320/2; 847/3; 881/1; 1317/2; 1500/3; 1653/2;

Query Match 4.7%; Score 103.5; DB 2; Length 3498;
Best Local Similarity 21.6%; Pred. No. 88;
Matches 79; Conservative 47; Mismatches 138; Indels 101; Gaps 17;

Qy 54 QVEAARK-AIRAAQVERVVPHERCCWCLCGCEVREHLSHGNLTVLYGGLLEHLASPEH 112
Db 2602 EMEAARKEAARAAEAYAAEQQR-----LELLRQEE 2634

Qy 113 KKATN-----KFWWENKAQVQMKKEFLVTPQDYARFKSMVKGLDSYEEKDKV--IKE 164
Db 2635 ERLRKEAERMRIQRENEERV- -QEQMRLEAEERIRRAEEERIQ--KELEDKVRREKE 2691

Qy 165 MAA-----QIREVQSROEVV--RSVLEP-----QAVPDPPEEGSSA 198
Db 2692 EAARQEKERQEARMEAREAEALSORMEQQRSSQONPYNQGGQYSQPPPPSYQSSY 2751

Qy 199 PRSMK-GMNSQVASSLQOPSNDLPPAPELDWMT-----GPSLTFIG 240
Db 2752 PNYQPGQGNQPNVQOPSHQSMQGHQAGYQOQTSNQMNMQOQNRQGGQGSFSG 2811

Qy 241 HQDI-----PGVG--NIHSGATPPMIQDEYIAGNQEIGPSYEFLEKEKOKLKKLPP 293
Db 2812 PGGINQPSQPGSYNGQGGQGGQGGQOQQRNPFNGQ-----QDMQPGGAALKMEAKP 2864

Qy 294 DRVGA-NFDHSSRTSAGWLPSPGFRLEXWTPLAVQTSPTKLKIQXRSSHIOKKANHALY 352
Db 2865 NEHAHQYQHTQNQLS--LAQKEKEQYFOAKNLQASQANAQQOQQRFGDV-VAGNVAGY 2921

Qy 353 QLPXG 357
Db 2922 GRPYG 2926

RESULT 12
T01083
hypothetical protein T10P11.5 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01083
R;Kaplan, N.; Johnson, D.; Schutz, K.; Gnoj, L.; Hoffman, J.; Till, S.; de la Bastide, M.
hi, M.; Martienssen, R.; Chen, E.Y.; Wilson, R.; McCombie, W.R.
submitted to the EMBL Data Library, November 1998
A;Description: Sequence of A. thaliana BAC T10P11 from chromosome IV.
A;Reference number: Z14248
A;Accession: T01083
A;Status: translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-3449 <KAP>
A;Cross-references: UNIPROT:Q9ZQX5; EMBL:AC002330; NID:g2262135; PID:g3892055
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 62/2; 86/3; 115/2; 138/3; 157/3; 198/3; 262/3; 284/3; 323/1; 726/3; 960/1; 99
A;Note: T10P11.5

Query Match 4.7%; Score 103; DB 2; Length 3449;
Best Local Similarity 20.1%; Pred. No. 94;
Matches 96; Conservative 46; Mismatches 143; Indels 192; Gaps 25;

Qy 6 WRRSRIVDPRAAXNSIPAHEGPFPCGCTGTAAAS----- 39
Db 2602 WNNDRATSTNAAS-----LHNSLSFGGKSSSTAVSVPIVNTDEKSETGSPKSSGKQWDE 2657

```
QY 40 -----TORQLKEAFERLL-POVEAARKAIRAAQOERYV--PEH-----ERCWCCL- 81
Db 2658 IKHVEESEKELKDDGEYLIRPYLEHLEKIRFRYNCRVVGLDKHDGIFLIGELCLYVIE 2717
QY 82 -----CCGCEVREHL-----HGNUTVLYGGLLEHLASPEHKKATNKW 120
Db 2718 NFYIDDHGICEKECEDELSTIDQAQGLKKQFHGSL-----ESKSKSSTLW 2763
QY 121 -----WENKAQVQMEKFLVT-----PODYARFKKSMVKGLDSYEE--KEDKVI 162
Db 2764 STTIKIGAVGGRAWAYGGGANGKEKRVVTGNLPHPMWK-----LDSVHILKRDYEL 2817
QY 163 KENAAQIREVE-----OSROEVRVSVL-----EPOAVDPPEGS-- 196
Db 2818 RRVAVEIFSMDCNDLLVPHKKEREVEFRNLLAMNLPNSMLDTTIGSAKQESKGSRL 2877
QY 197 -----SAPRSWKGMNSQVASSIQPSNL-----DLPAPELDWM-----ETGPSLTF 238
Db 2878 FKLMAKSFTKRWQ--NGEI-SNFQYLMHLNTLAGRGYSDLTQYPPVPMILADYDGESL-- 2932
QY 239 IGHQDIPGVGNIHSGATPPWMIQDEYIAGNOEIGPSYEEFLKEKEQ----- 286
Db 2933 -----DLSDNNFRK-LDKPMGQCTPE---GEEFRRKRYESWDDPEVPQPHYGSHYSSAGI 2984
QY 287 ----KLKKLPDRV-----GANFDHSSRTSAGWLPFGPRLEXWTPPLAVQSTPKLK 334
Db 2985 VLFYLIRLPFSAENQKLOGGQFDHADRL-----FNSIRETLWSAAGKGNSTSVK 3034

RESULT 13
S60248
protein-tyrosine kinase (EC 2.7.1.112) CAK-beta - human
N:Alternate names: cell adhesion kinase-beta
C:Species: Homo sapiens (man)
C>Date: 10-Apr-1996 #sequence revision 27-Feb-1997 #text_change 16-Aug-2004
C:Accession: S60248; G02330; B57434
R:Lev, S.; Moreno, H.; Martinez, R.; Canoll, P.; Peles, E.; Musacchio, J.M.; Plozman, G.; Nature 376, 737-745, 1995
A:Title: Protein tyrosine kinase PYK2 involved in Ca(2+)-induced regulation of ion channel
A:Reference number: S60248; MUID:95379967; PMID:7544443
A:Accession: S60248
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1009 <LEV>
A:Cross-references: UNIPROT:Q14289; EMBL:U33284; NID:g988304; PIDN:AAC50203.1; PID:g98830
R:Sasaki, T.; Nagura, K.; Sasaki, H.
submitted to the EMBL Data Library, December 1995
A:Reference number: H01067
A:Accession: G02330
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-22,'G',24-434,'L',436-1009 <SAS>
A:Cross-references: EMBL:U43522; NID:g1165218; PIDN:AAC05330.1; PID:g1165219
R:Sasaki, H.; Nagura, K.; Ishino, M.; Tobitoka, H.; Kotani, K.; Sasaki, T.
J. Biol. Chem. 270, 21206-21219, 1995
A:Title: Cloning and characterization of cell adhesion kinase beta, a novel protein-tyro
A:Reference number: A57434; MUID:95403356; PMID:7673154
A:Accession: B57434
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-22,'G',24-150 <SA2>
A:Cross-references: GB:D45853; NID:g1000676; PIDN:BAA08289.1; PID:d1008884; PID:g1000677
C:Genetics:
A:Gene: CAKBeta
C:Superfamily: protein kinase homology
C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
F:423-686/Domain: protein kinase homology <KIN>
F:431-439/Region: protein kinase ATP-binding motif

Query Match 4.7%; Score 102.5; DB:2; Length 1009;
Best Local Similarity 22.4%; Pred. No. 20;
Matches 84; Conservative 47; Mismatches 97; Indels 147; Gaps 23;
```

```
QY 119 FWNENKAQVQMEK-----FLVT-----PODYARFKKSMVKGLDSYEEK 157
Db 627 FWLENKDVIGVLEKGDRLPKPDLCPVLYTMTRCWDYDSDRPRFTSLVCSLSVDYQWE 686
QY 158 EDKVIKEMAAQIREVEQSRQREVRS--VLEPQAVDPDPEEGSSAPRSWKGMNSQVASSLQQ 215
Db 687 KDIA-----MEQERNARYRTPKILEFTAFQEPKPSRPK-----YRP 724
QY 216 PSNLDLPAPPELDWM-----ETGPSLTFIGHQDIPG-VGNIHSGATPP-----WM 259
Db 725 PPQTNL-LAPKLQFQVPEGLCASSPTLT--SPMEYPSFVNSLH--TPPLHRHNVPKRRHS 778
QY 260 IQDEYI--AGNOETGPSYE-----EFLKEKEKOKL-----KKLPPDRVGANF 300
Db 779 MREEDFIOPSSREAAQQLWEAEKVMRQILDKQKQWVEDYQWLROBEKSLDP-MVYWN- 836
QY 301 DHSRT---SAGWLPFGPRLEXWTPPLAVQSTPKLKIQXRSRSHIQKAN----- 348
Db 837 DKSLPTPEKEVGYLEFTGP-----PQKPRLGAQS-----IOPTANLDRTTDLVY 881
QY 349 -----HALYQLPXGXKPKSTKPLLYLPPKFFIIVFLRKQTYSFIXNKVLPFLGLRGL 401
Db 882 LNVMLVRAVLLEL-----KNELCQLPPEGYVVV-----KNVGLTLRKL 921
QY 402 SSL-----SIPSTSR 411
Db 922 GSVDDLPLPSLSSSR 936

RESULT 14
E86373
protein T23E23.10 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: E86373
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L. Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86373
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-762 <STO>
A:Cross-references: UNIPROT:Q9LRA0; GB:AE005172; NID:g9369408; PIDN:AAF87156.1; GSPDB:GN
C:Genetics:
A:Gene: T23E23.10
A:Map position: 1
C:Superfamily: Arabidopsis thaliana hypothetical protein F18N11.140

Query Match 4.7%; Score 102; DB:2; Length 762;
Best Local Similarity 19.5%; Pred. No. 15;
Matches 66; Conservative 49; Mismatches 99; Indels 124; Gaps 15;

QY 118 KFWNENKAQVQMEKFLVTQD-----YARFKKSMVKGLDSYEEK---BDKV 161
Db 456 KIFWRSLYQADIDGRRLETVDDVPVPEPEIPGTSILSFKEAMERGDKLTDLKLVMDSE 515
QY 162 IKEMAAQIREVEQSRQEVRSVLEPQAVDPPE-----GSSA---PRSWKGMNSQVASSLQ 214
Db 516 IKRCIVRVQGIE-----EYVADQLEKEAEKQYEDMRPGSAADFVPPSYQPRDTQSKALV- 570
QY 215 QPSNLDLPAPPELDWMETGPSLTFIGHQDIPGVGNIHSGATPPWMIODEYIAGNQIIGP 274
Db 571 -----IHSGRDPAPTTIESLSGSDS--- 592
QY 275 SYEEFLKEKEKQKLKKLPDRVGANFD-----HSSRT 306
```

```
Db      593  SDEEDKKKKKKQISKKRKHVGTLESGEKGRKPSKYKGEBTTECKQTRKQKSSS 652
Qy      307  SAGMLP-SFGPRLEXWTPLAVQTSTPKLKLOQXRSSHQKKAHALYQL----- 354
Db      653  SSSYRPOSSSARLQA-LPITATISSPSCQL-----VSSSDSSSSYQLIPLQNSSTSS 704
Qy      355  PXGKKPKSTKPLLYL---PPKFFIIVLRKQTYSFIXF 389
Db      705  PSGS--SSIDCIVHLRRPARLYL-----QQTTRFVQF 735
```

RESULT 15

```
T01275
hypothetical protein At2g19270 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F27F23.7
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: T01275; F84574
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, May 1998
A:Description: Arabidopsis thaliana chromosome II BAC F27F23 genomic sequence.
A:Reference number: Z14177
A:Accession: T01275
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-359 <ROU>
A:Cross-references: UNIPROT:O64560; EMBL:AC003058; NID:g3135250; PID:g3135257
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: F84574
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-359 <STO>
A:Cross-references: GB:AE002093; NID:g3135257; PIDN:AAC16457.1; GSPDB:GN00139
C:Genetics:
A:Gene: F27F23.7; At2g19270
A:Map position: 2
A:Introns: 305/3
```

```
Query Match          4.6%; Score 101; DB 2; Length 359;
Best Local Similarity 23.8%; Pred. No. 6.6;
Matches 50; Conservative 30; Mismatches 84; Indels 46; Gaps 8;

Qy      146  SMVKGLDSYERKDKVIKEMAAQIREVQSQEVVRSVLEPQAVDPPEGSAPRSWKGM 205
Db      3    SLVASAYSSDEEDLQPRREAFTVKSSGPSASSLFSALPQPKELKSSDGLGSSSRGK 62
Qy      206  NSQVASSLQOP-SNLDLPPAPELDWMETGPSLTFIGHDIP-GVGNIHSGATP-PWMIQD 262
Db      63  SSSFLSSLPPKPSISRKKNP-----NPS-----SIPRVVQIRLPVNPSPSLDD 108
Qy      263  EETIAGNQEIGPSYEEFLKEKEKOKLKLPPDRVGANFDHSRT-----SAGW 310
Db      109  EDD-----EEEEKARKKQKQMESASASHDSSVRSFLSAMPAPKSSQTIGA 153
Qy      311  LPSFGPRLEXWTPLAVQTSTPKLKLOQXRS 340
Db      154  LPSLGSGSGGRSRL--ETETPSIAPPQTDS 181
```

Search completed: June 7, 2005, 12:03:29
Job time : 17.7419 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:40:40 ; Search time 61.0244 Seconds
(without alignments)
3499.208 Million cell updates/sec

Title: US-09-155-676B-5

Perfect score: 2190

Sequence: 1 IGVTWRRSRIVDPRAXNS.....RGLESSLSIPSTSGRGRTHF 417

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1498	68.4	332	2 Q86UT8	Q86ut8 homo sapien
2	1302	59.5	338	2 Q7TNQ4	Q7tnq4 mus musculus
3	747	34.1	194	2 Q8C702	Q8c702 mus musculus
4	134	6.1	423	2 Q8W36	Q8w36 arabidopsis
5	118.5	5.4	1479	1 BAI8_MOUSE	Q22277 mus musculus
6	115.5	5.3	777	2 Q6GQ06	Q6gq06 xenopus lae
7	113	5.2	496	2 Q7ZUY4	Q7zuy4 brachydanio
8	112.5	5.1	487	2 Q8B6K0	Q8b6k0 soybean dwa
9	112.5	5.1	489	2 Q91QJ6	Q91qp6 soybean dwa
10	112.5	5.1	489	2 Q8B6M4	Q8b6m4 soybean dwa
11	112	5.1	1893	2 Q8CJ14	Q8cj14 rattus norv
12	111.5	5.1	850	1 STNA_DROME	Q24211 drosophila
13	111	5.1	536	2 Q8IH09	Q8ihd9 drosophila
14	111	5.1	539	2 Q8ML54	Q8mls4 drosophila
15	111	5.1	1266	2 Q7PCX6	Q7pcx6 macaca fasc
16	110.5	5.0	464	2 Q7QU17	Q7qui7 giardia lam
17	110.5	5.0	674	2 Q80W63	Q80w63 mus musculus
18	110.5	5.0	675	1 MTJF1_MOUSE	Q7243 mus musculus
19	110.5	5.0	675	2 Q8JWJ8	Q8jw8 mus musculus
20	110.5	5.0	1878	2 Q7LA35	Q7la35 mycoplasma
21	110	5.0	496	2 Q8P0V1	Q8p0v1 brachydanio
22	110	5.0	675	2 Q8BSY2	Q8bsy2 mus musculus
23	109.5	5.0	446	2 Q7XT92	Q7xt92 oryza sativ
24	109.5	5.0	489	2 Q8B6K2	Q8b6k2 soybean dwa
25	109.5	5.0	850	2 Q8XQ55	Q8xqes ralstonia s
26	109.5	5.0	1109	2 Q84HE0	Q84he0 oryza sativ
27	109.5	5.0	1109	2 Q7XFH5	Q7xfh5 oryza sativ
28	109	5.0	561	2 Q7XWH8	Q7xwh8 oryza sativ
29	109	5.0	1981	2 Q923Z7	Q923z7 rattus norv
30	109	5.0	5262	1 MLL2_HUMAN	O14686 homo sapien
31	108.5	5.0	1225	1 CTD2_HUMAN	Q9uqb3 homo sapien

ALIGNMENTS

RESULT 1

Q86UT8 PRELIMINARY; PRT; 332 AA.

AC Q86UT8; 01-JUN-2003 (Tremblrel. 24, Created)

DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)

DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)

DE DLNB14 protein.

GN Name=DLNB14;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Kubo T., Arai Y., Ohira M., Gamou T., Maeno G., Sakiyama T.,

RA Toyoda A., Hattori M., Sakaki Y., Nakagawara A., Ohki M.,

RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB094093; BAC76047.1; -

SQ SEQUENCE 332 AA; 37974 MW; 341209B75BE56349 CRC64;

Query Match 68.4%; Score 1498; DB 2; Length 332;

Best Local Similarity 89.1%; Pred. No. 4.3e-95;

Matches 286; Conservative 5; Mismatches 24; Indels 6; Gaps 2;

Qy 28 FFCRGFTTAASTORQLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87

Db 14 FFCRGHVYSRKQKQKLEALLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 73

Qy 88 REHLSHGNLTVLXGGLLEHLASPEHKATNKFWMENKAQVQMEKFLVTPQDYARFKKSM 147

Db 74 REHLSHGNLTVLXGGLLEHLASPEHKATNKFWMENKAQVQMEKFLVTPQDYARFKKSM 133

Qy 148 VKGLDSYEEKDKVKEMAAQIREVQSRQEVRSVLEPQAVPPPEGSSAPRSWKGNNS 207

Db 134 VKGLDSYEEKDKVKEMAAQIREVQSRQEVRSVLEPQAVPPPEGSSAPRSWKGNNS 193

Qy 208 QVASSLQOPSNLDLPPAPELDMMETGSLTFIGHQDIPGVGNTHSGATPPMMIODEEVIA 267

Db 194 QVASSLQOPSNLDLPPAPELDMMETGSLTFIGHQDIPGVGNTHSGATPPMMIODEEVIA 253

Qy 268 GNOEIGFSYEEFLKEKEKQKLLKPLPPRVGANGFHSRSTAGWLPSPFGRLXETWPLAVQ 327

Db 254 GNOEIGFSYEEFLKEKEKQKLLKPLPPRVGANGFHSRSTAGWLPSPFGRLXETWPLAVQ 310

Qy 328 TSTPKLKLQ---QXRSSHIQK 345

Db 311 QSRHQFKEAAAMKKQSHTEK 331

RESULT 2

Q7TNQ4

ID Q7TNQ4 PRELIMINARY; PRT; 338 AA.

AC Q7TNQ4;

[illegible]

Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A., Faney J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickens M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalek U., Smalhus D.E., Schnertch A., Schein J.E., Jones S.J., Marra M.A.;

"Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences";

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[2]

SEQUENCE FROM N.A.

RC TISSUE=Whole body;

RA STRAUSBERG R.;

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; ZDB-GENE-030826-14; nuc2a.

DR YFNI; BCD-GENE-030826-14; nuc2a.

DR GO; GO:0005509; F:calcium ion binding; IEA.

DR InterPro; IPR002048; EF-Hand.

DR InterPro; IPR010983; EF-Hand_like.

DR Pfam; PF00036; ehand; 2.

DR SMART; SM00054; EFh; 2.

DR PROSITE; PS00018; EF HAND; UNKNOWN 1.

DR PROSITE; 496 AA; 58136 MW; 4FD038C12A455AE2 CRC64;

SQ SEQUENCE

Query Match 5.2%; Score 113; DB 2; Length 496;

Best Local Similarity 20.8%; Pred. No. 17;

Matches 64; Conservative 47; Mismatches 94; Indels 102; Gaps 16

Qy 13 DPRAAXNSIAHCGPFCGRTGTTAASTQRLKEAFERLLPQVEAARKATRAAQVERVVP 72

Db 246 DKPTFFNLHDTNGDGF-----DEQELSLFTKELEKI-----YDP 281

Qy 73 -----EHERCCWCLCGCEVRHLSH---GNLTVLYGGLLEHLASPEHKKATNKF 119

Db 282 TNEEDDMVEEER-----LMRHRVNVNVDNDRLV-SLDFLVATKKKEFLPDP 332

Qy 120 WENKAEVQMKKFLVTPQDYARFKSMV-----KGLDSYEEKED--KVIKEMAAQI 169

Db 333 SWET-----LEQNAQYTEEMRFEFEQLVRQEEDLNQKAADLOKOREDRLERQQEQLNAQK 387

Qy 170 REVEQSRQEVVRSLVLEQAVDPDE--EGSAPRSWGMNSQVASSLQOPSNLDLPAPPE 226

Db 388 IEIQQAQVEHMER--IKTKQVPPFPSEILEGNAPVES-----VGQDQPPVP- 429

Qy 227 LDNMETGPSLTFIGHQDIPGVGNHSGATPPWMIQDEEYIAGNOIEGPSYEEFLKEKEQ 286

Db 430 LEHQPLPP-----GHQDVP-----PPAAQQHHDELQNALQGEHQNTPHNQ-- 472

Qy 287 KLKKLPP 293

Db 473 ---PLPP 476

RESULT 8

Q8B6K0 PRELIMINARY; PRT; 487 AA.

ID Q8B6K0

AC Q8B6K0

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Readthrough protein (Fragment).

GN Name=ORF5;

OS Soybean dwarf virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Luteoviridae;

OC Luteovirus.

OX NCBI_TaxID=12049;

RN [1]

```

Qy      227 LDWMETGPTSTFIQHQPDPGVNIHGATPPWMIQDEEYIAGNQEIIGPSVEEFLEKEKQ 286
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      430 LEHQPLP-----GHQDVP-----FPAAQQHHDELQNQQALGQHQTPHENQ-- 472
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      287 KLKKLPP 293
       |||||
Db      473 ---PLPP 476

RESULT 8
Q8B6KO PRELIMINARY; PRT; 487 AA.
AC Q8B6KO;
DT 01-MAR-2003 (TrEMBLrel_23, Created)
DT 01-MAR-2003 (TrEMBLrel_23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel_24, Last annotation update)
DE Readthrough protein (Fragment).
GN Name=ORF5;
OS Soybean dwarf virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Luteoviridae;
OC Luteovirus.
OX NCBI_TaxID=12049;
RN [1]

```

SEQUENCE FROM N.A.
 RC STRAIN=DS;
 RA Terauchi H., Honda K., Yamagishi N., Kanematsu S., Ishiguro K.,
 RA Hidaka S.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB076050; BAC54091.1; -;
 DR InterPro; IPR000893; Luteo ORF6.
 DR InterPro; IPR002929; PLRV ORF5.
 DR Pfam; PF01690; PLRV ORF5; 1.
 DR PRINTS; PR00910; LVIRUSORF6.
 FT NON_TER
 SQ SEQUENCE 487 AA; 54357 MW; 70F63DC6BE9B7B9A CRC64;

Query Match 5.1%; Score 112.5; DB 2; Length 487;
 Best Local Similarity 23.1%; Pred. No. 18;
 Matches 74; Conservative 49; Mismatches 100; Indels 97; Gaps 17;

QY 107 LASPEHKATNKF-----WENKAQVQMKKEFLVTPQDYARFKSMVKGLDSYEEKE 158
 Db 196 LVAPAVKK-TAKYNFCVSGDWTDRDMFQVSVVLDEHLEGARSSQYVRKTL-----RP 249
 QY 159 DKVIKEMAAQIREVQROEVRSVLEPQAVDPPEGGSSAPRSWKGMNSQVASSL-----213
 Db 250 GHISVNRSHRLQEMPTRED-----FSEEDNSSSDSNTYKTDSSPMTVHLAGDDE 299
 QY 214 ----QQPSNLDLPPAPPELDWMTGTPSLTFIGHQDI-----PGVGNL-----250
 Db 300 NYTTQSDDDVD-----PTLRDAEVMKFAQVGAQRFPALOGAKPIGPDRSIHRFKPSATL 355
 QY 251 ----HSGATPPWMIODEEYIAGNQIEIGPSYBEFL---KEKEQO-----KLKK-----LPPD 294
 Db 356 KSAFHSGLTP-----DELYNAGAEQADHDFEHFVSRGKEKLEQEDPMAKLREAEADLKPD 410
 QY 295 RVGANFDHSSRTSAGWLPFGRLXWTPLAVQTSTPKLKLOQXRSSHIQKANHLYQL 354
 Db 411 SDNEDNDY-DRTASVFNPDAG---DHWSGIKIK---PRLKQDQMRST-----L 451

QY 355 PXGX-KPKSTKPLLLYLPKPF 373
 Db 452 STGTLLKGGSLPKPKELPPRF 471

RESULT 9
 Q91QP6 PRELIMINARY; PRT; 489 AA.
 AC Q91QP6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE ORF5 protein (Fragment).
 GN Name=ORF5;
 OS Soybean dwarf virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Luteoviridae;
 OC Luteovirus.
 OC NCBI_TaxID=12049;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DP;
 RX MEDLINE=21578621; PubMed=11722012;
 RA Terauchi H., Kanematsu S., Honda K., Mikoshiba Y., Ishiguro K.,
 RA Hidaka S.;
 RT "Comparison of complete nucleotide sequences of genomic RNAs of four
 RT Soybean dwarf virus strains that differ in their vector specificity
 RT and symptom production."
 RL Arch. Virol. 146:1885-1898(2001).
 DR EMBL; AB038150; BAB62838.1; -;
 DR Pfam; PF01690; PLRV ORF5; 1.
 DR PRINTS; PR00910; LVIRUSORF6.
 FT NON_TER
 SQ SEQUENCE 489 AA; 54669 MW; EF81399B4BDBF747 CRC64;

Query Match 5.1%; Score 112.5; DB 2; Length 489;
 Best Local Similarity 23.8%; Pred. No. 18;

Matches 62; Conservative 37; Mismatches 70; Indels 91; Gaps 15;
 QY 178 EVRVSVLEP-----QAVP-----DPEGSSAPRSWKGMNSQVA-----210
 Db 241 QYVRKTLRPGHISVNRSHRLQEMPPMRDFSEEDNSSSDSNTYKTDNSGPMTVHFVKAGDK 300
 QY 211 SSLQPSNLDLPPAP-ELDWMETGSLTFIGHQDI-----PGVGNL-----250
 Db 301 ESYTTQSDDDVEPSPHDADVMKFS---GFVGAKDRFVLQGADPIPGDGRSVRRFKPSATL 357
 QY 251 ----HSGATPPWMIODEEYIAGNQIEIGPSYBEFL-KEKEK-----QKLKLLPPD 294
 Db 358 KSAFHSGLTP-----DELYNAGAEQADHDFEHFVSRGKEKLEQEDPMAKLREAEADLKPD 412
 QY 295 RVGANFDHSSRTSAGWLPFGRLXWTPLAVQTSTPKLKLOQXRSSHIQKANHLYQL 354
 Db 413 SDNEDNDY-DRTAPVFNPDAG---DPWSGIKIK---PRLKQDQMRST-----L 453

QY 355 PXGX-KPKSTKPLLLYLPKPF 373
 Db 454 STGTLLKGGSLPKPKELPPRF 473

RESULT 10
 Q8B6M4 PRELIMINARY; PRT; 489 AA.
 AC Q8B6M4;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Readthrough protein (Fragment).
 GN Name=ORF5;
 OS Soybean dwarf virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Luteoviridae;
 OC Luteovirus.
 OC NCBI_TaxID=12049;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DS;
 RA Terauchi H., Honda K., Yamagishi N., Kanematsu S., Ishiguro K.,
 RA Hidaka S.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB076038; BAC54067.1; -;
 DR InterPro; IPR000893; Luteo ORF6.
 DR InterPro; IPR002929; PLRV ORF5.
 DR Pfam; PF01690; PLRV ORF5; 1.
 DR PRINTS; PR00910; LVIRUSORF6.
 FT NON_TER
 SQ SEQUENCE 489 AA; 54634 MW; FA4E6F4900DBEA77 CRC64;

Query Match 5.1%; Score 112.5; DB 2; Length 489;
 Best Local Similarity 21.9%; Pred. No. 18;
 Matches 70; Conservative 54; Mismatches 101; Indels 95; Gaps 16;

QY 107 LASPEHKATNKF-----WENKAQVQMKKEFLVTPQDYARFKSMVKGLDSYEEKE 158
 Db 196 LVAPAVKK-TAKYNFCVSGDWTDRDMFQVSVVLDEHLEGARSSQYVRKTL-----RP 249
 QY 159 DKVIKEMAAQIREVQROEVRSVLEPQAVDPPEGGSSAPRSWKGMNSQVASSLQQPSN 218
 Db 250 GHISVNRSHRLQEMPTRED-----FSEEDNSSSDSNTYKTDNSRMTVHVKAGD 299
 QY 219 LDLPAPPELDWMTGTPSLT-----FIGHQD---IPGVGNL-----250
 Db 300 EESYATQTDQDVE--PTLLDADVMKFGVGFVARDRFTLQGAEPDGTSSRRFRPSATL 357
 QY 251 ----HSGATPPWMIODEEYIAGNQIEIGPSYBEFL-KEKEK-----QKLKLLPPD 294
 Db 358 KSAFHSGLTP-----DELYNAGAEQADHDFEHFVSRGKEKLEHGEDPMAKLREAEADLKPD 412
 QY 295 RVGANFDHSSRTSAGWLPFGRLXWTPLAVQTSTPKLKLOQXRSSHIQKANHLYQL 354
 Db 413 SDNEDNDY-DRTASVFNPDAG---DPWSGIKIK---PRLKQDQMRST-----L 453

```
QY 355 PXGK-KPKSTKPLLYLPKPKF 373
DB 454 STGTGLGSLGPKPKELPPRF 473

RESULT 11
Q8CJ14
ID Q8CJ14 PRELIMINARY; PRT; 1893 AA.
AC Q8CJ14;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Nestin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=22573441; PubMed=12686602;
RA Chou Y.H., Khoun S., Herrmann H., Goldman R.D.;
RT "Nestin promotes the phosphorylation-dependent disassembly of vimentin intermediate filaments during mitosis.";
RL Mol. Biol. Cell 14:1468-1478(2003).
RN [2]
SEQUENCE FROM N.A.
RA Chou Y.-H., Herrmann H., Goldman R.D.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the intermediate filament family.
DR EMBL; AF538924; AAN33053.1; -.
DR HSSP; P08670; 1GK7.
DR GO; GO:0005882; C:intermediate filament; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR011000; Apolp_III_like.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; Filament; 1.
DR PROSITE; PS00226; IF; 1.
DR KW Intermediate filament.
SQ SEQUENCE 1893 AA; 208796 MW; 03AE6B616A1A7623 CRC64;

Query Match 5.1%; Score 112; DB 2; Length 1893;
Best Local Similarity 20.6%; Pred. NO. 1.le-02;
Matches 73; Conservative 53; Mismatches 131; Indels 98; Gaps 14;

QY 35 FTAASQ--ROLKEAFERLLPQ--VEAARKAIRAAQVERVYVEHERCCWCLCCGVEVREH 90
DB 663 FPGAEQDQMLERLVEKEDQSFPSPEEDQEAQCRPLQKENQEP-----LGVEEAE- 711
QY 91 LSHGNLTLYGLLEHLASPEHKATKNFWMENKAEVQMEKFLVTPQDY-ARFKKSMVK 149
DB 712 -----QILLERLLEKESQES-----LRSPEEDQEAQGRSLQK 743
QY 150 GLD---SYEKEDKVI-----KENAAQIREVQSQEVRVRLVBPQAVDPPEEGSSAPRS 201
DB 744 QNQPGLGVEEAQGQILERLIEKESQESLSRAEEDQEAQCRSL-----QKENQEPGLG 794
QY 202 WKGNWSQVASSIQPSNLDLPAPELDWMETGPSLTFIGHODIPGVGNHSGATPPWMIQ 261
DB 795 YEEAEQDLERLIEKESQESLSRAEEDQEAQGRSLQ-----K 831
QY 262 DEVIAGNQETGPSVEEFLKEKQKLLPPD--RVGANFDHSRSTAGWLPSPGPRLE 319
DB 832 ENQEPGLVEEAEDQMLERLIEKESQESLSRAEEDQEAQGRSLYLER---NQE 888
QY 320 XWTPPLAVQTSFPLKVL-----QQRSSHIQKANHLYQLPXPQKPKSTKPLLYL 369
DB 889 TFVPLESRNORPLRSLEVEEQRIVKPLKVSQDSLSGL----AEENVQLRVL 939

RESULT 12
STNA_DROME
ID STNA_DROME STANDARD; PRT; 850 AA.

AC Q24211; Q9WSM8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Stoned A protein (Stoneda) (Stn-A).
GN Name=stnA; ORFNames=CG40306;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
STRAIN=Oregon-R; TISSUE=CNS;
RX MEDLINE=97001127; PubMed=8844157;
RA Andrews J., Smith M., Merakovsky J., Coulson M., Hannan F.,
RA Kelly L.E.;
RT "The stoned locus of Drosophila melanogaster produces a dicistronic transcript and encodes two distinct polypeptides.";
RL Genetics 143:1699-1711(1996).
RN [2]
SEQUENCE FROM N.A.
STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E.C., Spadling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou S., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [3]
GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
```


systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [4]
 CC SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=22426071; PubMed=12537574;
 RA Hoskins R.A., Smith C.D., Carlson J.W., Carvalho A.B., Halpern A.,
 Kaminker J.S., Kennedy C., Mungall C.J., Sullivan B.A., Sutton G.G.,
 Yasuhara J.C., Wakimoto B.T., Myers E.W., Celniker S.E., Rubin G.M.,
 Karpen G.H.;
 RA "Heterochromatic sequences in a Drosophila whole-genome shotgun
 assembly.";
 RT Genome Biol. 3:RESEARCH0085.1-RESEARCH0085.16(2002).
 [5]
 CC FUNCTION.
 RP MEDLINE=99042043; PubMed=9822725;
 RA Stimson D.T., Estes P.S., Smith M., Kelly L.E., Ramaswami M.;
 RT "A product of the Drosophila stoned locus regulates neurotransmitter
 release.";
 RL J. Neurosci. 18:9638-9649(1998).
 [6]
 CC FUNCTION. SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.
 RP MEDLINE=99337940; PubMed=10407025;
 RA Fergestad T., Davis W.S., Broadie K.;
 RT "The stoned proteins regulate synaptic vesicle recycling in the
 presynaptic terminal.";
 RL J. Neurosci. 19:5847-5860(1999).
 [7]
 CC INTERACTION WITH SYT, AND MUTANT STN-TS2.
 RP MEDLINE=20524362; PubMed=11069931;
 RA Phillips A.M., Smith M., Ramaswami M., Kelly L.E.;
 RT "The products of the Drosophila stoned locus interact with synaptic
 vesicles via synaptotagmin.";
 RL J. Neurosci. 20:8254-8261(2000).
 [8]
 CC FUNCTION, AND SUBCELLULAR LOCATION.
 RP MEDLINE=21114085; PubMed=11160392;
 RA Fergestad T., Broadie K.;
 RT "Interaction of stoned and synaptotagmin in synaptic vesicle
 endocytosis.";
 RL J. Neurosci. 21:1218-1227(2001).
 [9]
 CC FUNCTION.
 RP MEDLINE=21212245; PubMed=11312288;
 RA Stimson D.T., Estes P.S., Rao S., Krishnan K.S., Kelly L.E.,
 Ramaswami M.;
 RT "Drosophila stoned proteins regulate the rate and fidelity of synaptic
 vesicle internalization.";
 RL J. Neurosci. 21:3034-3044(2001).
 CC -!- FUNCTION: Adapter protein involved in endocytic recycling of
 synaptic vesicles membranes. May act by mediating the retrieval of
 synaptotagmin protein Syt from the plasma membrane, thereby
 facilitating the internalization of multiple synaptic vesicles
 from the plasma membrane.
 CC -!- SUBUNIT: Interacts with the second C2 domain of Syt.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; colocalizes with synaptic
 vesicle pools. Colocalizes with the endocytic network within
 synaptic boutons.
 CC -!- DEVELOPMENTAL STAGE: Present at synaptic connections both in the
 CNS and in neuromuscular junctions in the mature embryo (20-22h)
 and throughout larval development. In the third instar larva, it
 is expressed in all synaptic boutons types, including I, II and
 III boutons.
 CC -!- DOMAIN: The Asp-Pro-Phe (DPF) motifs, which are found in many
 presynaptic proteins, are thought to mediate an interaction with
 Alpha-adaptin (By similarity).
 CC -!- MISCELLANEOUS: StnB, which is involved in the same pathway, is
 derived from the same dicistronic transcript that encodes these
 two different proteins.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC
 DR EMBL; U54982; AAC16665.1; --
 DR PIR; T13352; T13352; stnA.
 DR FlyBase; FBgn0016976; stnA.
 DR GO; GO:0030139; C:endocytic vesicle; IDA.
 DR GO; GO:0005886; C:plasma membrane; IDA.
 DR GO; GO:0008021; C:synaptic vesicle; IGI.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR GO; GO:0008099; P:synaptic vesicle endocytosis; IMP.
 KW Endocytosis; Repeat; Synapse.
 FT DOMAIN 26 290 Interaction with Syt.
 FT DOMAIN 9 17 Poly-Lys.
 FT DOMAIN 72 77 Poly-Ser.
 FT DOMAIN 344 350 Poly-Glu.
 FT DOMAIN 431 434 Poly-Pro.
 FT SITE 224 226 DPF 1.
 FT SITE 461 463 DPF 2.
 FT SITE 535 537 DPF 3.
 FT SITE 666 668 DPF 4.
 FT SITE 755 757 DPF 5.
 FT SITE 823 825 DPF 6.
 FT MUTAGEN 35 K->M: In stn-TS2; induce behavioral
 FT defects. Homozygous adults are viable but
 FT exhibit sluggish and uncoordinated
 FT behavior.
 FT F -> L (in Ref. 1).
 FT CONFLICT 463 463
 FT SEQUENCE 850 AA; 93108 MW; 2773741640B73757 CRC64;
 SQ
 Query Match 5.1%; Score 111.5; DB 1; Length 850;
 Best Local Similarity 23.2%; Pred. No. 44;
 Matches 89; Conservative 52; Mismatches 151; Indels 91; Gaps 21;
 QY 38 ASTORQLKEAFERLI--POVEAA-RKATRAAQVERVYVPEHERCCWCLCCGCEVREHLHG 94
 DB 270 ANPKELRGKIQNLLLSLSIELADSEALLAATSNAEFQHNLL-----DDLDELS 322
 QY 95 NLTVLYGGLLEHLASPEHKATNKFWMENKAQVQMKELVTPDYARFKSMVKGLDSY 154
 DB 323 SVPIDLSVSL-HLHLIKHKQPVEE---EELEQKRENGQLLNP-DLSEF-----DSL 369
 QY 155 EEKEDKVIKEMAAQIREVQSRQVRSVLEPQVPPPEEGSSAPRSWKGNVSQVASSLO 214
 DB 370 KDEEDFEPAELAAE-SLTKKEEVTVSQVLPVLA-QLPTEAFEA-GSWAEFEQ---SGQ 423
 QY 215 QPSNLDLPPAPELDWMETGSLTFIGHQDIPGVGNHSGATPPWMIQDEEYIAGNQEIGP 274
 DB 424 EPGKPKRPPPPVRP--PTGPHI-----VPGA-----IYSEDEENP 458
 QY 275 SYEFLKEKEKQKLKLPDRVGANFDHSSRTSAGWLPSP-GRLEKWTPLAVOTS----- 329
 DB 459 EDDPFTNTYAEQVIKKTIVLEEDDDFPRAEEHATEPPFLAAPQORDLAGSATDLSQVVP 518
 QY 330 ---TPKLKLOQXRSRH--IQKANHALYQLPGKKPKSTKLLLYLPKFFIIVLRKQTY 384
 DB 519 APLAFTLSVDOEAEDDFDFTSVAVALVQ-----PKSTE-----LRFLERELL 561
 QY 385 SFIXFNKVLFFGLRG--LESSLS 405
 DB 562 N-----YSGLDGVTLKHLSLS 576
 RESULT 13
 QBIHD9
 ID Q8IHD9 PRELIMINARY; PRT; 536 AA.
 AC Q8IHD9;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE AT11052p.

GN Name=CG30412;CG30416; ORFNames=CG30416;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT001295; AA01050.1; -
DR FlyBase; FBgn0050416; CG30416.
SQ SEQUENCE 536 AA; 60593 MW; C5B462E6E254C0E CRC64;

Query Match 5.1%; Score 111; DB 2; Length 536;
Best Local Similarity 29.7%; Pred. No. 26;
Matches 33; Conservative 21; Mismatches 39; Indels 18; Gaps 4;

Qy 152 DSYEEKEDKVIKEMAAQIREVEQSRQVRSVLE---PQAVDPDEEGSSAPRSKMGNSQ 208
Db 244 ESKDAKEEIVQYEYKTERDFDDPKAKVEKTIKVIKIPVDSPEAESENSSPPSGIPPPSQ 303

Qy 209 VASSLQOPSNL--DLPPAPELDMMETGSLTFIGHQDIPGVGNHSGATPP 257
Db 304 LPSDIPPPSQLSDIPPPSQL-----PS-----DIPPPSQLPSGTTPP 341

RESULT 14
Q8MLS4
ID Q8MLS4 PRELIMINARY; PRT; 539 AA.
AC Q8MLS4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG30416-PA.
GN ORFNames=CG30416;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yeaman M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Fowler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M., P.G.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J., Ye J.,
RA Williams S.M., Woodruff, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Swirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Swirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomic perspective";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003450; AAM68240.2; -
DR FlyBase; FBgn0050416; CG30416.
SQ SEQUENCE 539 AA; 60907 MW; 03526863C861005C CRC64;

Query Match 5.1%; Score 111; DB 2; Length 539;
Best Local Similarity 29.7%; Pred. No. 27;
Matches 33; Conservative 21; Mismatches 39; Indels 18; Gaps 4;

Qy 152 DSYEEKEDKVIKEMAAQIREVEQSRQVRSVLE---PQAVDPDEEGSSAPRSKMGNSQ 208
Db 244 ESKDAKEEIVQYEYKTERDFDDPKAKVEKTIKVIKIPVDSPEAESENSSPPSGIPPPSQ 303

Qy 209 VASSLQOPSNL--DLPPAPELDMMETGSLTFIGHQDIPGVGNHSGATPP 257
Db 304 LPSDIPPPSQLSDIPPPSQL-----PS-----DIPPPSQLPSGTTPP 341

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw_model

Run on: June 7, 2005, 11:57:41 ; Search time 20.7652 Seconds
(without alignments)
2390.553 Million cell updates/sec

Title: US-09-155-676B-5
Perfect score: 2390
Sequence: 1 IGVTWRRRSRIVDPRAAXNS.....RGLESSLSTPSTGRGTRHF 417

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 557933 seqs, 119041758 residues

Total number of hits satisfying chosen parameters: 557933

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*

- 1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pcp.*
- 2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pcp.*
- 3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pcp.*
- 4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pcp.*
- 5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pcp.*
- 6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pcp.*
- 7: /cgn2_6/ptodata/1/paa/US11_NEW_COMB.pcp.*
- 8: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2178	99.5	417	5	US-09-155-676B-5
2	670	30.6	156	6	US-10-450-763-49939
3	554.5	25.3	741	6	US-10-450-763-49940
4	503	23.0	312	6	US-10-450-763-44039
5	124.5	5.7	567	6	US-10-703-032-105976
6	111.5	5.1	850	7	US-11-097-143-32898
7	111	5.1	838	7	US-11-097-143-20796
8	108.5	5.0	830	6	US-10-450-763-38504
9	108.5	5.0	1225	6	US-10-940-774A-6063
10	108.5	5.0	1233	6	US-10-450-763-35355
11	107.5	4.9	1540	6	US-10-940-774A-7037
12	106.5	4.9	605	6	US-10-703-032-111745
13	105.5	4.8	2246	6	US-10-450-763-36209
14	105	4.8	669	7	US-11-097-143-35838
15	103.5	4.7	753	6	US-10-940-774A-6676
16	103.5	4.7	3111	7	US-11-097-143-7773
17	103.5	4.7	3224	6	US-10-528-183-8
18	103	4.7	844	6	US-10-450-763-36592
19	102.5	4.7	667	6	US-10-940-774A-7759
20	102.5	4.7	777	6	US-10-940-774A-8706
21	102.5	4.7	777	6	US-10-940-774A-8707
22	102.5	4.7	1014	6	US-10-940-774A-11533
23	102.5	4.7	1096	1	PCT-US03-10870-747
24	102	4.7	1448	1	PCT-US03-10870-998
25	101.5	4.6	968	7	US-11-097-143-11865

ALIGNMENTS

RESULT 1

US-09-155-676B-5

Sequence 5, Application US/09155676B

GENERAL INFORMATION: David

APPLICANT: WALLACH, David

APPLICANT: MALININ, Nikolai

APPLICANT: BOLDIN, Mark

APPLICANT: KOVALENKO, Andrei

APPLICANT: METT, Igor

TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR

FILE OF INVENTION: PREPARATION AND USE

FILE REFERENCE: WALLACH-21

CURRENT APPLICATION NUMBER: US/09/155.676B

PRIOR FILING DATE: 1999-01-04

PRIOR APPLICATION NUMBER: PCT/IL97/00117

PRIOR FILING DATE: 1997-04-01

PRIOR APPLICATION NUMBER: IL 117800

PRIOR FILING DATE: 1996-04-02

PRIOR APPLICATION NUMBER: IL 119133

PRIOR FILING DATE: 1996-08-26

NUMBER OF SEQ ID NOS: 22

SOFTWARE: PatentIn version 3.3

SEQ ID NO 5

LENGTH: 417

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (18)..(18)

OTHER INFORMATION: Xaa can be any naturally occurring amino acid

FEATURE:

NAME/KEY: misc feature

LOCATION: (320)..(320)

OTHER INFORMATION: Xaa can be any naturally occurring amino acid

FEATURE:

NAME/KEY: misc feature

LOCATION: (338)..(338)

OTHER INFORMATION: Xaa can be any naturally occurring amino acid

FEATURE:

NAME/KEY: misc feature

LOCATION: (356)..(356)

OTHER INFORMATION: Xaa can be any naturally occurring amino acid

FEATURE:

NAME/KEY: misc feature

LOCATION: (358)..(358)

OTHER INFORMATION: Xaa can be any naturally occurring amino acid

FEATURE:

NAME/KEY: misc feature

; LOCATION: (388)...(388)		; OTHER INFORMATION: Xaa can be any naturally occurring amino acid	
US-09-155-676B-5			
Query Match		99.5%; Score 2178; DB 5; Length 417;	
Best Local Similarity		100.0%; Pred. No. 1.9e-155;	
Matches 417; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	IGVTRRSRIVDPRRAAXNSIRAHGPPFCGRGTTAASTQRLKEAFERLLPQVEAARK	60
Db	1	IGVTRRSRIVDPRRAAXNSIRAHGPPFCGRGTTAASTQRLKEAFERLLPQVEAARK	60
QY	61	ATRAAQERYVPEHERCCWCLCCGCEVREHLSHGNTLVLYGGLLHLASPEHKATNKF	120
Db	61	ATRAAQERYVPEHERCCWCLCCGCEVREHLSHGNTLVLYGGLLHLASPEHKATNKF	120
QY	121	WENKAQVQKEFLVTPQDYARFKSMVKGLDSYEKEDKVIKEMAAQIREVEQSRQ	180
Db	121	WENKAQVQKEFLVTPQDYARFKSMVKGLDSYEKEDKVIKEMAAQIREVEQSRQ	180
QY	181	RSVLPEQAVDPDEEGSSAPRSWKGMNSQVASSLQPSNLDLPPAPELDMETGPSLTIFG	240
Db	181	RSVLPEQAVDPDEEGSSAPRSWKGMNSQVASSLQPSNLDLPPAPELDMETGPSLTIFG	240
QY	241	HODIPGVGNHSGATPPMWIQDEEYIAGNQEIGPSYEEFLKEKEKQKLLPDRVGANF	300
Db	241	HODIPGVGNHSGATPPMWIQDEEYIAGNQEIGPSYEEFLKEKEKQKLLPDRVGANF	300
QY	301	DHSSRTSAGWLPSFGPRLEXWTPAVQTSTPKLKLQXRSASHIQKANHALYQLPXGKXP	360
Db	301	DHSSRTSAGWLPSFGPRLEXWTPAVQTSTPKLKLQXRSASHIQKANHALYQLPXGKXP	360
QY	361	KSTKPLLYLPKPFIIIVFLRKQTSYFIKXNKVLPFGLRGLESLSIPSTSRGGRTHF	417
Db	361	KSTKPLLYLPKPFIIIVFLRKQTSYFIKXNKVLPFGLRGLESLSIPSTSRGGRTHF	417
RESULT 2			
US-10-450-763-49939			
; Sequence 49939, Application US/10450763			
; GENERAL INFORMATION:			
; APPLICANT: Hyseq, Inc			
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES			
; FILE REFERENCE: 790CIP3/US			
; CURRENT APPLICATION NUMBER: US/10/450,763			
; CURRENT FILING DATE: 2003-06-11			
; PRIOR APPLICATION NUMBER: PCT/US01/08631			
; PRIOR FILING DATE: 2001-03-30			
; PRIOR APPLICATION NUMBER: 09/540,217			
; PRIOR FILING DATE: 2000-03-31			
; PRIOR APPLICATION NUMBER: 09/649,167			
; PRIOR FILING DATE: 2000-08-23			
; NUMBER OF SEQ ID NOS: 60736			
; SOFTWARE: Custom			
; SEQ ID NO 49939			
; LENGTH: 156			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (1)...(156)			
; OTHER INFORMATION: Xaa = X or * as defined in Table 2			
US-10-450-763-49939			
Query Match		30.6%; Score 670; DB 6; Length 156;	
Best Local Similarity		77.9%; Pred. No. 8.2e-43;	
Matches 134; Conservative		4; Mismatches 12; Indels 22; Gaps 4;	
QY	6	WRRSRIVDPRRAAXNSIRAHGPPFCGRGTTAASTQRLKEAFERLLPQVEAARKAIRAA	65
Db	5	WSRRL-OPQAPA-----AAGGFG--GAAPAGA-----EVEAARKAIRAA	42
QY	66	QERYVPEHERCCWCLCCGCEVREHLSHGNTLVLYGGLLHLASPEHKATNKFWMENKA	125

```
; Sequence 44039, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 7901P3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 44039
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-450-763-44039

Query Match      23.0%; Score 503; DB 6; Length 312;
Best Local Similarity 49.3%; Pred. No. 6.7e-30;
Matches 113; Conservative 14; Mismatches 38; Indels 64; Gaps 6;

Qy 40 TORQLKEAFERL-----LPQVEAARKAIRAAQVRYVPEHERCCWCLCCGCEVREHL 91
Db 108 TQRLPAAADMPAAITGITDVEAARKAIRAAQVRYVPEHERCCWCLCCGCEVREHL 167
Qy 92 SHGNLTVLYGLLEHLASPEHKATNKFWMENKAEVQMKFVLTPDYARF----- 143
Db 168 SHGNLTVLYGLLEHLASPEHKATNKFWMENKAEVQMKFVLTPDYARSHLYGTPMG 227
Qy 144 ----KXSVKGL-----DSYEKEDKVIKEMAAQIREVSQSRQVRSV-LEPQA 188
Db 228 SAVPSQACLEALLIIVPGACVQSEES-----PAGSKDEPGEQVELKEEA 276
Qy 189 VPDPEEGSAPRSWKMGNSQVASSLQPPSNLDLPPAPELDMETGPSLT 237
Db 277 EAPVEDGSQP-----PPEPKDATPGEKAT 303

RESULT 5
US-10-703-032-105976
; Sequence 105976, Application US/10703032
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Andersen, Scott E.
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Cao, Yongwei
; APPLICANT: Masucci, James D.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53374)B
; CURRENT APPLICATION NUMBER: US/10/703,032
; CURRENT FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: 10/020,338
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 211164
; SEQ ID NO 105976
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(567)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_TA_394.pap
; US-10-703-032-105976

Query Match      5.7%; Score 124.5; DB 6; Length 567;
Best Local Similarity 23.9%; Pred. No. 0.37;
Matches 71; Conservative 46; Mismatches 99; Indels 81; Gaps 16;

Qy 35 FTAASTORQLKEAFERLLPQVEAARKAIRAAQVRYVPEHERCCWCLCCGCEVREHLSHG 94
Db 116 FNAVERPKEREDLFENYVLELQKKEKAKAAEHRRTAEYR----- 156
Qy 95 NLTVLYGLLEHLASPEHKATNKFWMENKAEVQMKFVLTPDYARFK-----S 146
Db 157 -----EFLSCDFIKA-NTQWRKVQDRLEDDER-----YARLEKIDRLDVFQD 198
Qy 147 MVKGLDSYEKEDKVIKEMAAQIREVSQSRQVRSVLEPQAVDPDPREGS-SAPRSWKGM 205
Db 199 YIRHLEKEEKEQKIRKE---QLRRQERKRDPERKMEEHVA---DGTJNAKTYWRDY 251
Qy 206 NSQVASS---LQPPSNLDLPPAPEL---DMETGPSLTFIGHODIPGV---GNIHSGATP-- 256
Db 252 CSQIKDSRAYLAVASNLSGMPKELFDDVMEELDK----QYDDRALIKDEVKSGKIPML 307
Qy 257 -PMWI-----QDEEYI-AGNQEIGPSYE---EFLKEKEKQKLKLPDPDRVGANF 300
Db 308 ASWTLDFQAAVTEDEKYKGVSNINIKLIYEDQIERLKEKDLKEAKK--RQRLGDNF 362

RESULT 6
US-11-097-143-32898
; Sequence 32898, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32898
; LENGTH: 850
; TYPE: PRT
; ORGANISM: DROSOPHILA
; US-11-097-143-32898

Query Match      5.1%; Score 111.5; DB 7; Length 850;
Best Local Similarity 23.2%; Pred. No. 5.8;
Matches 89; Conservative 52; Mismatches 151; Indels 91; Gaps 21;

Qy 38 ASTORQLKEAFERLL---PQVEAA-RKAIRAAQVRYVPEHERCCWCLCCGCEVREHLSHG 94
Db 270 ANPKRLKRGIQNLLLSIESIELADSEAEELAAATSNAPQHNL-----DDLDEELSES 322
Qy 95 NLTVLYGLLEHLASPEHKATNKFWMENKAEVQMKFVLTPDYARFKSMYKGLDSY 154
Db 323 SVPIDLSVSL-HLHLIKHKQFVEE---EELEQKRGRENQLNLP--DLSEF-----DSL 369
```


[illegible]

Db 223 SSOIAQHQKQKTVVKEDEGRRESINDRARRSPKLPKLSKKGERKWAPPKFLPHKYDVK 282
QY 223 -----PAPELDMMETGPSLT-----FIGHQDIPGVGNHSHGATPPMILQDEEYIA 267
Db 283 LQNEDKLIISNVPADSLIRTPPNKEIVRYFIHNAL-----RAGTGENAPVWVEDE--LV 336
QY 268 GNGEIGPSYEEFLKEKEK-----QKLKCLPPDR---VGANFDHSSRTSAGWLPSPFGP 316
Db 337 KKYSLPSKSFDFLLDPYKMYTLNPSKTKRNTGSPDKPSKSKTNDSSLS-----PLNP 391
QY 317 RLEKXW-----TPPLAVQTS-----TPKUKLQO-----XRSISHQKXANHAL 351
Db 392 KL--WCHVHLKKSLSGSPKLVKNSKNSKSPHEHLEMMKQMSPNKLIHTNFHIPKKG----445
QY 352 YQLPXCKXP--KSTKPL 366
Db 446 ---PPAKKPKGKSDKPL 459

RESULT 12
US-10-703-032-111745
; Sequence 111745, Application US/10703032
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Andersen, Scott E.
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Cao, Yongwei
; APPLICANT: Masucci, James D.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(53374)B
; CURRENT APPLICATION NUMBER: US/10/703,032
; CURRENT FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: 10/020,338
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 211164
; SEQ ID NO 111745
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_TA_6163.pep
US-10-703-032-111745

Query Match 4.9%; Score 106.5; DB 6; Length 605;
Best Local Similarity 21.5%; Pred. No. 8.9;
Matches 85; Conservative 53; Mismatches 132; Indels 125; Gaps 20;
QY 20 SIRAHEGPPFCGRGT--FTAASTQRL-----KEAFERLLPOVEAARK--AIRAQVERY 70
Db 93 SSRVNDGICDCGSEEGYNSCTCKNTWEAGAKARELKKVATYKSGVIRNOVEK- 151
QY 71 VPEHERCCWLCGCGVEHLSHGNLTVLYG-----GLLEHASPHEKKATNKFWNK 124
Db 152 -----AELATAKDEABLTCLKGEBKILQGLVDKL--KEQKRLIEKAESEER 195
QY 125 AEVQMEKEFLVTPQDYARFKSMVKGK-----PSYEKEKDKVKEMAAQIREVQ 174
Db 196 LIKEEERKRIKEVENKQRLRTIADPAQAEADSDTNEKAQDESKVVEHHDGDTIDHN 255
QY 175 SRQEVRSV-----LEPOAVPDPPEGSGAPRS-----WKG 204
Db 256 HSPGETSVVEVPDIKAGTCDDBPVTSAVPTTEODPTSVNSEGLSKAELGRLVASRTG 315
QY 205 MN-SQVASSLOQ--PSNLDLP--PAPELDMMETGPSLTTFIGHQDIP-----GVGNHSGAT 255
Db 316 ENVDVVGSKDKKGHEDELDPPEAE-----AFEDHDIPEPVEENVAGVHS---362
QY 256 PPMILQDEYIAGNOEI--GPSVEEFLKEKEKQKLK-KLPDRVGFANFDHSSRTSAG---W 310
Db 363 ---ETEDRRHKFEDELSNVSDDEYVDHDEPDSYKSDDDRKGD--DHSDLTASGQASW 417

QY 311 LPSFGPRLEXWTPLAVQTSPTPKLKLQXRSSHIQK 345
Db 418 LD-----KIQTIVQNVLQK 431
RESULT 13
US-10-450-763-36209
; Sequence 36209, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 36209
; LENGTH: 2246
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (434)..(449)
; OTHER INFORMATION: POLLEN ALLERGEN POA PI SIGNATURE domain identified by
; OTHER INFORMATION: EMATRIX, accession number PR00833H, p-value=3.077e-09, raw score
; OTHER INFORMATION: 2.30
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (4)..(95)
; OTHER INFORMATION: Gag P30 core shell protein domain identified by Pfam,
; OTHER INFORMATION: accession name Gag_P30, E-value=1.4e-27, Pfam score of 100.2
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2246)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-36209

Query Match 4.8%; Score 105.5; DB 6; Length 2246;
Best Local Similarity 22.1%; Pred. No. 56;
Matches 98; Conservative 64; Mismatches 135; Indels 147; Gaps 23;
QY 8 RSRIVDPRAAXNS----IRAHEGPPFCGRGTFTAASTQRLKEAFERLLPOVEAARKAIR 63
Db 1419 RELVKEPARAGDRQRPWELEEQQGRQFCS---TQAALQAMEREA--EQMGNELERLAALM 1473
QY 64 AAQVERVYPEHERCCWLCGCGVEHLSHGNLTVLYGGLLEHASPHEKKATNKFWNEN 123
Db 1474 ESQGOQ---QBER-----GQBER---VARLTQERGRAQADLAL---EKAA-----1510
QY 124 KAEVQMEKEFLVTPQ--DYARFKSMVKGKLSYEEKEDKVIKEM---AAQIREVEQSROE 178
Db 1511 RAELEMLQNALNEQRVEFATLQELAHALTEKSGKQDELAKRLGLEAAQKLEBELQOT 1570
QY 179 VVR-----SVLEPQAVDPDPPEGSSA 198
Db 1571 VKQLKEQLAKKEKHAHSGSGAQSEAAGRTPTGPKLEALRAEVSKLEQCCQKQOQADSL 1630
QY 199 PRSKWGNMSQVASSLQPSNLD-LPPAPELDMMETGPSLTTFIGH-----QDIP 245
Db 1631 ERS---LEAERASRAERDSDALETTLQGLEEKAQELGHSQSALASQAQRELAARFTKVQD--1685
QY 246 GVGNIHSGATPPMILQDEYIAGNOE-----IGPSVEEFL-----KEKEKQKLK 290
Db 1686 -----HSAEDEVKAQ---VARGQEAERKNLSLSSEEEVSIINRQVLEKEGESKELKR 1737
QY 291 LPPDRVGFANFDHSSRTSAGWLPSPGRLEXWTPLAVQTSPTPKLKLQXRSSHIQKXANHA 350

Db 1738 L-----VMAESKSOKLE-----ERLLQLQAETASNSARAAE-RSSALRELEHT 1781
Qy 351 YQLPXPXKKPKSTKPL--YLPPK 372
Db 1782 -----STQALVSELLPAK 1794

RESULT 14
US-11-097-143-35838
; Sequence 35838, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35838
; LENGTH: 669
; TYPE: PRT
; ORGANISM: DROSOPHILA

US-11-097-143-35838
Query Match 4.8%; Score 105; DB 7; Length 669;
Best Local Similarity 25.1%; Pred. No. 13;
Matches 65; Conservative 37; Mismatches 99; Indels 58; Gaps 11;
Qy 122 ENKAEVQMKKFLVTPDYARFKSMVKGDSYEKEDKVIKEMAAQIR-EVEQSRQEV- 179
Db 407 EDLEMEIK-----HERKELKSHGASTDSMEEDNENETMTVABEHHSDGEVSNGEVP 459
Qy 180 -----VRSVLEPQAVDPPEGSSAPRSWKGNMNOVASLQOPS---NLDLPP-APELDWM 230
Db 460 IEERPTTSSEKPSASELPEEGNSAPPALKKDVKELOAQAQAVSHAVNLLAPPKATEAE-- 517
Qy 231 ETGSLFTIGHQDIPGVGNHSGATPPMMDIEEYIAGNQEIGPSYBEFLKEKEKQKX 290
Db 518 -----PRTLRSKLSPOPP--VVDKK-----SAKOKKXKKKKQKQPOE 551
Qy 291 LPPDRVG--ANFDHSSRTSAGWL-----PSFGPRLEXWTP-----LAVQTSTPKLKLQX 338
Db 552 ASPLKSSDEENHGHRIRTNAGYTVVDEPTKVIILITSSGCMVRVEPTCKQKYFPR 611
Qy 339 RSSHIQKANHLYQLPXG 357
Db 612 AAATRKCMGSGARQDPPG 630

RESULT 15
US-10-940-774A-6676
; Sequence 6676, Application US/10940774A
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/10/940,774A
; CURRENT FILING DATE: 2004-09-15
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6676
; LENGTH: 753
; TYPE: PRT
; ORGANISM: Human
; US-10-940-774A-6676

Query Match 4.7%; Score 103.5; DB 6; Length 753;
Best Local Similarity 22.0%; Pred. No. 20;
Matches 84; Conservative 48; Mismatches 149; Indels 101; Gaps 19;
Qy 19 NSIRAH-----EGPFPCG-RGTTAASTQRLK-----EAFERLLPOVE 56
Db 185 HSLRIHVHVHTKKEPFCDVQCEKAFNTLYRLKAHQRLHTGKTFCNCESECGSKYFTTLS 244
Qy 57 AAKAIPRAAOVER-----YVPEH--ERCWCCLCCCE----- 86
Db 245 DLRRKHRTHTGKPPFRCDDHCGKFAAASHHLKTHVTHTGERTFFCFPSNGCEKTFSTQY 304
Qy 87 -VREHL-SHGNLTLYLGGILLEHLASPEHKK-----ATNKFWMENKAEVQMKKFL 134
Db 305 SLKSHMKGNDKSHYNALPQHGSEDTNHSCLSDLSLSTSELRENSSTTQGDULT 364
Qy 135 VTPQDYARFKSMVKGDSYEKEDKVIKEMAAQIREVQSRQEVRSVLEPQAVDPPEE 194
Db 365 ISP---AIIFESM-----FQNSDDTAIQEDPQQTASLTESFNGDAESVSD---VP-PST 411
Qy 195 GSSAPRS-----WKGMNSQVASSLQOPSNLDLPPAPELDWMETGPTSLTFIHQDIPGVGN 249
Db 412 GNSASLSLPLVLQPLGLSEPPQPLLPASAPSAPPAPSL-----GP-----GSQQ-AAFGN 460
Qy 250 IHSGATPPMMDIEEY-IAGNQEIGPSYBEFLKEKEKQKLLKLPDRVGFANFDHSSRTSA 308
Db 461 -----PPALLQPEVDPVPHSTOFAANHQEFLEPHQAPQ-PIVPGLSVVAGASASAAVA 513
Qy 309 GWLPSPGPRLEXWTP-LAVQTST 330
Db 514 SAVAAPAPPQSTTEPLPAMVQT 535

Search completed: June 7, 2005, 12:21:18
Job time : 22.7652 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: June 7, 2005, 11:52:21 ; Search time 186.04 Seconds
(without alignments)
2618.037 Million cell updates/sec

Title: US-09-155-676B-5
Perfect score: 2190
Sequence: 1 IGTRWRRSRIVDPRAAXNS.....RGLESSLSIFSTSGRGTRHF 417

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA.Main.*
1: /cgn2_6/ptodata/1/paa/PTCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US105_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
33: /cgn2_6/ptodata/1/paa/US107_COMB.pep.*
34: /cgn2_6/ptodata/1/paa/US108_COMB.pep.*
35: /cgn2_6/ptodata/1/paa/US109_COMB.pep.*
36: /cgn2_6/ptodata/1/paa/US110_COMB.pep.*
37: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2178	99.5	417	15	US-09-155-676-5 Sequence 5, Appli
2	2178	99.5	417	15	US-09-155-676A-5 Sequence 5, Appli
3	972.5	44.4	211	27	US-10-170-2058-36800 Sequence 9118, Ap
4	972.5	44.4	212	37	US-60-453-050-9118 Sequence 9118, Ap
5	972.5	44.4	212	37	US-60-453-135-9118 Sequence 9118, Ap
6	972.5	44.4	212	37	US-60-466-412-9118 Sequence 9118, Ap
7	795	36.3	1115	24	US-09-915-543-17 Sequence 17, Appl
8	795	36.3	1115	29	US-10-322-579-17 Sequence 17, Appl
9	795	36.3	1115	22	US-10-564-859-17 Sequence 17, Appl
10	731.5	33.4	216	22	US-09-758-463-1029 Sequence 1029, Ap
11	731.5	33.4	216	28	US-10-217-607-1029 Sequence 1029, Ap
12	670	30.6	156	1	PCT-US01-08631-49939 Sequence 49939, A
13	554.5	25.3	741	1	PCT-US01-08631-49940 Sequence 49940, A
14	517	23.6	189	1	PCT-US04-12047-1081 Sequence 1081, Ap
15	503	23.0	312	1	PCT-US01-08631-44039 Sequence 44039, A
16	309.5	14.1	65	1	PCT-US01-01354-17236 Sequence 17236, A
17	309.5	14.1	65	22	US-09-764-909-17236 Sequence 17236, A
18	309.5	14.1	65	26	US-10-092-399-17236 Sequence 17236, A
19	198	9.0	104	1	PCT-US98-23829-2 Sequence 2, Appli
20	198	9.0	104	15	US-09-185-936-2 Sequence 2, Appli
21	198	9.0	104	21	US-09-746-783-206 Sequence 206, App
22	198	9.0	104	32	US-10-689-742-206 Sequence 206, App
23	183	8.4	71	17	US-09-330-337-292 Sequence 292, App
24	148	6.8	389	30	US-10-424-599-209186 Sequence 209186, A
25	134	6.1	423	21	US-09-708-427-26860 Sequence 26860, A
26	117.5	5.4	443	37	US-60-556-841-5931 Sequence 5931, Ap
27	117.5	5.4	1040	30	US-10-425-115-308861 Sequence 308861, A
28	116	5.3	367	28	US-10-219-999-39253 Sequence 39253, A
29	116	5.3	367	30	US-10-425-114-41619 Sequence 41619, A
30	116	5.3	367	30	US-10-425-114A-41619 Sequence 41619, A
31	116	5.3	367	37	US-60-312-544-8751 Sequence 8751, Ap
32	116	5.3	368	37	US-60-324-109-26189 Sequence 26189, A
33	116	5.3	369	28	US-10-219-999-59765 Sequence 59765, A
34	116	5.3	369	30	US-10-425-114-69227 Sequence 69227, A
35	116	5.3	369	30	US-10-425-114A-69227 Sequence 69227, A
36	116	5.3	653	28	US-10-219-999-40981 Sequence 40981, A
37	116	5.3	653	30	US-10-425-114-51821 Sequence 51821, A
38	116	5.3	653	30	US-10-425-114A-51821 Sequence 51821, A
39	116	5.3	653	37	US-60-324-109-17151 Sequence 17151, A
40	112	5.1	124	28	US-10-219-999-49996 Sequence 49996, A
41	112	5.1	124	28	US-10-219-999-50911 Sequence 50911, A
42	112	5.1	160	30	US-10-425-115-219898 Sequence 219898, A
43	112	5.1	253	30	US-10-425-114-62537 Sequence 62537, A
44	112	5.1	253	30	US-10-425-114-67550 Sequence 67550, A
45	112	5.1	253	30	US-10-425-114A-62537 Sequence 62537, A

ALIGNMENTS

RESULT 1
US-09-155-676-5
; Sequence 5, Application US/09155676
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
; FACTOR (TRAF), THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/155,676
FILING DATE: 04-JAN-1999
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IL97/00117
FILING DATE: 01-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 117800
FILING DATE: 02-APR-1996
APPLICATION NUMBER: IL 119133
FILING DATE: 26-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH=21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 417 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-155-676-5

Query Match 99.5%; Score 2178; DB 15; Length 417;
Best Local Similarity 100.0%; Pred. No. 2.1e-186;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGVTRWRSRIVDPRAAXNSIRAHGPPFCGRGTFTAASTQOLKEAFERLLPQVEAARK 60
DB 1 IGVTRWRSRIVDPRAAXNSIRAHGPPFCGRGTFTAASTQOLKEAFERLLPQVEAARK 60
QY 61 AIRAAQVERVYVPEHERCCWCLCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNPF 120
DB 61 AIRAAQVERVYVPEHERCCWCLCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNPF 120
QY 121 WENKAEVQMKFVLTPQDYARFKKSMVKGLDSYEKEDKVIKEMAAQIREVEQSQREV 180
DB 121 WENKAEVQMKFVLTPQDYARFKKSMVKGLDSYEKEDKVIKEMAAQIREVEQSQREV 180
QY 181 RSVLEPQAVPDPPEGSSAPRSWKGMNSQVASSLQOPSNLDLPPAPELDMETGPSLTF 240
DB 181 RSVLEPQAVPDPPEGSSAPRSWKGMNSQVASSLQOPSNLDLPPAPELDMETGPSLTF 240
QY 241 HQDIFGVGNHSGATPPWMIQDEEYIAGNQIEGSPSYEEFLKEKEQKLLKLPDRVGANF 300
DB 241 HQDIFGVGNHSGATPPWMIQDEEYIAGNQIEGSPSYEEFLKEKEQKLLKLPDRVGANF 300
QY 301 DHSSTTSAGWLPSPGRLXWTPLAVQTSPTPKLQQRSSHIQKANHLYQLPXGKXP 360
DB 301 DHSSTTSAGWLPSPGRLXWTPLAVQTSPTPKLQQRSSHIQKANHLYQLPXGKXP 360
QY 361 KSTKPLLYLPKPFVILVRKQTSYIFXFNKVLFFGLRGLESSLSPTSRGGRTHF 417
DB 361 KSTKPLLYLPKPFVILVRKQTSYIFXFNKVLFFGLRGLESSLSPTSRGGRTHF 417

RESULT 2

US-09-155-676A-5
Sequence 5, Application US/09155676A
GENERAL INFORMATION:
APPLICANT: WALLACH, David
APPLICANT: MALININ, Nikolai

APPLICANT: BOLDIN, Mark
APPLICANT: KOVALENKO, Andrei
APPLICANT: METT, Igor
TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
STREET: 624 Ninth Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/155,676A
FILING DATE: 04-JAN-1999
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IL97/00117
FILING DATE: 01-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 117800
FILING DATE: 02-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 119133
FILING DATE: 26-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH=21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 417 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-155-676A-5

Query Match 99.5%; Score 2178; DB 15; Length 417;
Best Local Similarity 100.0%; Pred. No. 2.1e-186;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGVTRWRSRIVDPRAAXNSIRAHGPPFCGRGTFTAASTQOLKEAFERLLPQVEAARK 60
DB 1 IGVTRWRSRIVDPRAAXNSIRAHGPPFCGRGTFTAASTQOLKEAFERLLPQVEAARK 60
QY 61 AIRAAQVERVYVPEHERCCWCLCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNPF 120
DB 61 AIRAAQVERVYVPEHERCCWCLCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNPF 120
QY 121 WENKAEVQMKFVLTPQDYARFKKSMVKGLDSYEKEDKVIKEMAAQIREVEQSQREV 180
DB 121 WENKAEVQMKFVLTPQDYARFKKSMVKGLDSYEKEDKVIKEMAAQIREVEQSQREV 180
QY 181 RSVLEPQAVPDPPEGSSAPRSWKGMNSQVASSLQOPSNLDLPPAPELDMETGPSLTF 240
DB 181 RSVLEPQAVPDPPEGSSAPRSWKGMNSQVASSLQOPSNLDLPPAPELDMETGPSLTF 240
QY 241 HQDIFGVGNHSGATPPWMIQDEEYIAGNQIEGSPSYEEFLKEKEQKLLKLPDRVGANF 300
DB 241 HQDIFGVGNHSGATPPWMIQDEEYIAGNQIEGSPSYEEFLKEKEQKLLKLPDRVGANF 300
QY 301 DHSSTTSAGWLPSPGRLXWTPLAVQTSPTPKLQQRSSHIQKANHLYQLPXGKXP 360
DB 301 DHSSTTSAGWLPSPGRLXWTPLAVQTSPTPKLQQRSSHIQKANHLYQLPXGKXP 360

Db 301 DHSRTSAGWLPSPGRLEXTPTPLAVQTSPTKLKQQVRSRSHIOKKANHLYQLPXGKXP 360
Qy 361 KSTRPLLYLPKFFIIVFLRKQTSYFIFXFNKVLFFGLRGLSSLSIPSTSRGGRTHF 417
Db 361 KSTRPLLYLPKFFIIVFLRKQTSYFIFXFNKVLFFGLRGLSSLSIPSTSRGGRTHF 417

RESULT 3
US-10-170-205E-36800
; Sequence 36800, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36800
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-36800

Query Match 44.4%; Score 972.5; DB 27; Length 211;
Best Local Similarity 74.5%; Pred. No. 2.5e-78;
Matches 190; Conservative 1; Mismatches 7; Indels 57; Gaps 1;

Qy 28 FFCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87
Db 14 PFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 73

Qy 88 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKFLVTPDYARFKKSM 147
Db 74 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKFLVTPDYARFKKSM 133

Qy 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGMNS 207
Db 134 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGMNS 133

Qy 208 QVASSLOQPSNLDLPPAPELDMETGPSLTFIGHQDIPGVGNIHSGATPPPMIODEEYIA 267
Db 172 -----DIPGVGNIHSGATPPPMIODEEYIA 196

Qy 268 GNQEIGPSYEEFLKE 282
Db 197 GNQEIGPSYEEFLKE 211

US-10-170-205E-36800

Query Match 44.4%; Score 972.5; DB 27; Length 211;
Best Local Similarity 74.5%; Pred. No. 2.5e-78;
Matches 190; Conservative 1; Mismatches 7; Indels 57; Gaps 1;

Qy 28 FFCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87
Db 14 PFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 73

Qy 88 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKFLVTPDYARFKKSM 147
Db 74 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKFLVTPDYARFKKSM 133

Qy 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGMNS 207
Db 134 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGMNS 171

Qy 208 QVASSLOQPSNLDLPPAPELDMETGPSLTFIGHQDIPGVGNIHSGATPPPMIODEEYIA 267
Db 172 -----DIPGVGNIHSGATPPPMIODEEYIA 196

Qy 268 GNQEIGPSYEEFLKE 282
Db 197 GNQEIGPSYEEFLKE 211

US-10-170-205E-36800

Query Match 44.4%; Score 972.5; DB 27; Length 212;
Best Local Similarity 74.5%; Pred. No. 2.5e-78;
Matches 190; Conservative 1; Mismatches 7; Indels 57; Gaps 1;

Qy 28 FFCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87
Db 14 PFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 73

Qy 88 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKFLVTPDYARFKKSM 147
Db 74 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKFLVTPDYARFKKSM 133

Qy 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGMNS 207
Db 134 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGMNS 171

Qy 208 QVASSLOQPSNLDLPPAPELDMETGPSLTFIGHQDIPGVGNIHSGATPPPMIODEEYIA 267
Db 172 -----DIPGVGNIHSGATPPPMIODEEYIA 196

Qy 268 GNQEIGPSYEEFLKE 282
Db 197 GNQEIGPSYEEFLKE 211

US-10-170-205E-36800

Query Match 44.4%; Score 972.5; DB 37; Length 212;
Best Local Similarity 74.5%; Pred. No. 2.5e-78;
Matches 190; Conservative 1; Mismatches 7; Indels 57; Gaps 1;

Qy 28 FFCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87
Db 14 PFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 73

Qy 88 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKFLVTPDYARFKKSM 147
Db 74 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKFLVTPDYARFKKSM 133

Qy 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGMNS 207
Db 134 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGMNS 171

Qy 208 QVASSLOQPSNLDLPPAPELDMETGPSLTFIGHQDIPGVGNIHSGATPPPMIODEEYIA 267
Db 172 -----DIPGVGNIHSGATPPPMIODEEYIA 196

Qy 268 GNQEIGPSYEEFLKE 282
Db 197 GNQEIGPSYEEFLKE 211

US-60-453-135-9118

RESULT 5
US-60-453-135-9118
; Sequence 9118, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9118
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(212)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-453-135-9118

Query Match 44.4%; Score 972.5; DB 37; Length 212;
Best Local Similarity 74.5%; Pred. No. 2.5e-78;
Matches 190; Conservative 1; Mismatches 7; Indels 57; Gaps 1;

Qy 28 FFCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87
Db 14 PFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 73

Qy 88 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKFLVTPDYARFKKSM 147
Db 74 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKFLVTPDYARFKKSM 133

Qy 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGMNS 207
Db 134 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGMNS 171

Qy 208 QVASSLOQPSNLDLPPAPELDMETGPSLTFIGHQDIPGVGNIHSGATPPPMIODEEYIA 267
Db 172 -----DIPGVGNIHSGATPPPMIODEEYIA 196

Qy 268 GNQEIGPSYEEFLKE 282
Db 197 GNQEIGPSYEEFLKE 211

US-60-453-135-9118


```

RESULT 6
US-60-466-412-9118
; Sequence 9118, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9118
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(212)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-466-412-9118

```

Query Match	44.4%;	Score 972.5;	DB 37;	Length 212;
Best Local Similarity	74.5%;	Pred. No. 2.5e-78;		
Matches 190;	Conservative 1;	Mismatches 7;	Indels 57;	Gaps 1
Qy	28	FFCGRGTFTAASQTQQLKEAFERLLPQVEAAARKAIRAAQOVERVYVPEHERCCWCLCGCEV	87	
Db	14	FFCGRHVVYRKHQRLKEALERLLPQVEAAARKAIRAAQOVERVYVPEHERCCWCLCGCEV	73	
Qy	88	REHLSHGNIITVLYGGLLEHLASPEHKATKATKFWENKAEVOMKEKFLVTPQDYARPKKM	147	
Db	74	REHLSHGNIITVLYGGLLEHLASPEHKATKATKFWENKAEVOMKEKFLVTPQDYARPKKM	133	
Qy	148	VKGLDSYEEKEDKVIKEMAAQIREVESOROSVRSLVLEPQAVDPDEGSSAPRSWKGMNS	207	
Db	134	VKGLDSYEEKEDKVIKEMAAQIREVESQREVVRSLVLE-----	171	
Qy	208	QVASSLQOPSNLDDLPAPPELDMWTGTSPLTFIGHQDIPGVGNTHSGATPPWMIQDEEYIA	267	
Db	172	-----DIPGVGNTHSGATPPWMIQDEEYIA 196		
Qy	268	GNQEICPSYEFLKE	282	
Db	197	GNQEICPSYEFLKE	211	

```

RESULT 7
US-09-915-543-17
; Sequence 17, Application US/09915543
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich
; APPLICANT: FROESCH, Barbara
; APPLICANT: KRAMPS, Thomas
; APPLICANT: PETER, Oliver
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY
; FILE REFERENCE: Q60361
; CURRENT APPLICATION NUMBER: US/09/915,543
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,502
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Human lgs-1
US-09-915-543-17

```

Query Match 36.3%; Score 795; DB 24; Length 1115;

	Best Local Similarity	55.6%, Pred. No. 2.9e-61,			
	Matches	180; Conservative	17; Mismatches	67; Indels	60; Gaps
					8;
<hr/>					
QY	28	FFCGRGTTAASTORQLKEAFERLLPQVEAARAKAIRAAQVRYVPYHERCCWCLCCGCEV	87		
		:			
Db	828	FFCGRGHVYSRKHQRQLKEALERLLPQVEAARAKAIRAAQVRYVPYHERCCWCLCCGCEV	887		
		:			
QY	88	REHLSHGNLTVLTYGGLELHLELASPEHKATNKFWMENKAEOVMKKEFLVTQDYARFKKSM	147		
		:			
Db	888	REHLSHGNLTVLTYGGLELHLELASPEHKATNKFWMENKAEOVMKKEFLVTQDYARFKKSM	947		
		:			
QY	148	VKGLDSEYEEDKVIKEMAAQIREVESQREVRSVLEPQAVPDPEGSSAPRSWKGMNS	207		
		:			
Db	948	VKGLDSEYEEDKVIKEMAAQIREVESQREVRSVLE-----TGPPR-----	990		
		:			
QY	208	QVASLOOPSNLDDLPPAPELDMMETGSLTFIGHQDIPGVGNITHSGATPPMMIODEBYIA	267		
		: : : :	:	:	:
Db	991	-YALTVRSPAVLSR-----RTLKSGAFPQP--TPEAHPQ	1021		
		:			
QY	268	GNOEIGFSEYEELKEKEQKLK-KLPDRVCANFDHSRSTAGWLPSFGRLXKWTPPLAV	326		
		: : :			
Db	1022	ARCLCAPRGALKPDPPTVKLGVPHTTRKARPAAKTISP-----RPRCRQAQ---	1072		
		:			
QY	327	QTSTPKLKL---QQXRSSHIOKKA	347		
		: :			
Db	1073	-NKTSQSLQAGARKAKTAHLLOTKA	1095		
		: :			

RESULT 8
US-10-322-579-17
; Sequence 17, Application US/10322579
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich
; APPLICANT: FROESCH, Barbara
; APPLICANT: KRAMPS, Thomas
; APPLICANT: PETER, Oliver
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
; FILE REFERENCE: Q60361
; CURRENT APPLICATION NUMBER: US/10/322,579
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US/09/915,543
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,502
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Human lgs-1
US-10-322-579-17

Query Match	36.3%	Score 795;	DB 29;	Length 1115;
Best Local Similarity	55.6%;	Pred. No. 2.9e-61;		
Matches 180;	Conservative 17;	Mismatches 67;	Indels 60;	Gaps 8;
Qy	28	FFCGRGTTAAS ¹ QROLKEAFERLLPQVEAAR ² KAIRAAQ ³ VERV ⁴ PPERCCWLCGCC ⁵ CEV	87	
Db	828	FFCGRGHVYSRKHQR ¹ KEALELLPQVEAAR ² KAIRAAQ ³ VERV ⁴ PPERCCWLCGCC ⁵ CEV	887	
Qy	88	REHLSHGNTLVLYGGLLEHLASPEHKATNKFWENKA ¹ EVQMEKEFLVTPQDYARFKSM	147	
Db	888	REHLSHGNTLVLYGGLLEHLASPEHKATNKFWENKA ¹ EVQMEKEFLVTPQDYARFKSM	947	
Qy	148	VKGLDSYEKEDKVIKEMAAQ ¹ IREVQSQREVSV ² SL ³ EPQAVDPDPEEGSAPPSWKGMS	207	
Db	948	VKGLDSYEKEDKVIKEMAAQ ¹ IREVQSQREVSV ² SL ³ EPQAVDPDPEEGSAPPSWKGMS	990	
Qy	208	QVASSLQQPSNLN ¹ DLPPAPELDMMETG ² PSLFTTGHQDIPGVGNTHSGATPPMWIQDBEYIA	267	
Db	991	-YALT ¹ VRSPAVLSR-----RTUKSGAPPO--TPRAHPO	1021	

```
QY 268 GNOEIGSYEBFLKEKEKQK-LKLPDRVGANFDHSRTSAGWLPFGPPLXKWTPLAV 326
DB 1022 ARCLCAPRGALKPEPPGRTKLGVPHHTTKRPHAAKTSF-----RPRCTRQAP--- 1072
QY 327 QTSTPKLKL---QXRSISHOKKA 347
DB 1073 -NKTQSLQLAGKARKATLHLQTKA 1095

RESULT 9
US-10-664-859-17
; Sequence 17, Application US/10664859
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich
; APPLICANT: FROESCH, Barbara
; APPLICANT: KRAMPS, Thomas
; APPLICANT: PETER, Oliver
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY
; FILE REFERENCE: Q60361
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
; CURRENT APPLICATION NUMBER: US/10/664,859
; PRIOR FILING DATE: 2003-09-22
; PRIOR APPLICATION NUMBER: US/09/915,543
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,502
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Human lgs-1
US-10-664-859-17

Query Match 36.3%; Score 795; DB 32; Length 1115;
Best Local Similarity 55.6%; Pred. No. 2.9e-61;
Matches 180; Conservative 17; Mismatches 67; Indels 60; Gaps 8;

QY 28 FFCGRGFTTAATORQLKEAFERLLPOVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87
DB 828 FFCGRGHVYSKQHQKQKLEALLPOVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 887
QY 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWENKAQVOMKEKFLVTPDYARFKKSM 147
DB 888 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWENKAQVOMKEKFLVTPDYARFKKSM 947
QY 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGN 207
DB 948 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLE-----TGPPR----- 990
QY 208 QVASSLOQPSNLDLPAPPELDMETGSLTFIGHQDIPGVNIIHSGATPPHMIODESYIA 267
DB 991 -YALTVRSPAVLSR-----RTLKSGAPFPQ--TPEAHPQ 1021
QY 268 GNOEIGSYEBFLKEKEKQK-LKLPDRVGANFDHSRTSAGWLPFGPPLXKWTPLAV 326
DB 1022 ARCLCAPRGALKPEPPGRTKLGVPHHTTKRPHAAKTSF-----RPRCTRQAP--- 1072
QY 327 QTSTPKLKL---QXRSISHOKKA 347
DB 1073 -NKTQSLQLAGKARKATLHLQTKA 1095

RESULT 10
US-09-758-463-1029
; Sequence 1029, Application US/09758463
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM011
; CURRENT APPLICATION NUMBER: US/09/758,463
```

```
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1029
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-463-1029

Query Match 33.4%; Score 731.5; DB 22; Length 216;
Best Local Similarity 78.8%; Pred. No. 1.3e-56;
Matches 141; Conservative 3; Mismatches 12; Indels 23; Gaps 1;

QY 28 FFCGRGFTTAATORQLKEAFERLLPOVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87
DB 61 FFCGRGHVYSKQHQKQKLEALLPOVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 120
QY 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWENKAQVOMKEKFLVTPDYARFKKSM 147
DB 121 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWENKAQVOMKEKFLVTPDYARFKKSM 180
QY 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGN 206
DB 181 VKGLDSYEEDKVKIKEMAAQIR-----EWSRADRRWFGLS 216

RESULT 11
US-10-217-607-1029
; Sequence 1029, Application US/10217607
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM011CIN
; CURRENT APPLICATION NUMBER: US/10/217,607
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 09/758,463
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1029
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-217-607-1029

Query Match 33.4%; Score 731.5; DB 28; Length 216;
Best Local Similarity 78.8%; Pred. No. 1.3e-56;
Matches 141; Conservative 3; Mismatches 12; Indels 23; Gaps 1;

QY 28 FFCGRGFTTAATORQLKEAFERLLPOVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87
DB 61 FFCGRGHVYSKQHQKQKLEALLPOVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 120
QY 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWENKAQVOMKEKFLVTPDYARFKKSM 147
DB 121 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWENKAQVOMKEKFLVTPDYARFKKSM 180
QY 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGN 206
DB 181 VKGLDSYEEDKVKIKEMAAQIR-----EWSRADRRWFGLS 216

RESULT 12
PCT-US01-08631-49939
; Sequence 49939, Application PC/TUS0108631
```

```

; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 49939
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)-(156)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-49939

Query Match 30.6%; Score 670; DB 1; Length 156;
Best Local Similarity 77.9%; Pred. No. 2.7e-51;
Matches 134; Conservative 4; Mismatches 12; Indels 22; Gaps 4;

QY 6 WRSRIVDRAAXNSIRAHEGPFPCGRTFTTAASTORQLKEAFERLLLPQVEAARKAIRAA 65
DB 5 WSRARL-QQAPAA---AAEGGF--GEAPAPGA-----EVEAARKAIRAA 42

QY 66 QERYVPERHCWCWCLCCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNKFWWENKA 125
DB 43 QERYVPERHCWCWCLCCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNKFWWENKA 102

QY 126 EVQMKKEFLVTPQDYARFKKSMVKGGLDSYEEDKVKIKEMAAQIREVEQSRQ 177
DB 103 EVQMKKEFLVTPQDYARFKKSMVKGGLDSYEEDKVKIKEMAAQIREVEQSRQ 154

RESULT 13
PCT-US01-08631-49940
; Sequence 49940, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 49940
; LENGTH: 741
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: DOMAIN
; LOCATION: (632)...(650)
; OTHER INFORMATION: BURKITT'S LYMPHOMA RECEPTOR SIGNATURE domain identified by
; OTHER INFORMATION: EMATRIX, accession number PR00564E, p-value=1.918e-23, raw score
; OTHER INFORMATION: 12.64
; NAME/KEY: DOMAIN
; LOCATION: (502)...(736)
; OTHER INFORMATION: 7 transmembrane receptor domain identified by Pfam, accession
; OTHER INFORMATION: name 7tm_1, E-value=5.2e-81, Pfam score of 259.5
PCT-US01-08631-49940

Query Match 25.3%; Score 554.5; DB 1; Length 741;
Best Local Similarity 40.7%; Pred. No. 7e-40;
Matches 135; Conservative 18; Mismatches 52; Indels 127; Gaps 11;

```

```

QY 66 QERYVPERHCWCWCLCCGCEVREHLSHGNTLVLYGGLLEHL-----ASPEHKK----- 114
DB 175 KERYVPERHCWCWCLCCGCEVREHLSHGNTLVLYGGLLEHLARTASIVAAPNWKQFRCL 234
QY 115 ATNKFWMENKAEVQMKKEFLVTPQDYARFKKSMVKGGLDSYEEDKVKIKEMAAQIREVEQ 174
DB 235 STGK--WINNG-----IVFQEDRI-----IDTYDTDES-----QK 263
QY 175 SRQEVRSVLEPQAVPDPEEGSSAPRSWKGMNSQVASSLQOPSNLDLPPAPELDWMETGP 234
DB 264 SR----- 265
QY 235 SLTFIGHQDIPGVGNIHSGATPPWMIQDEEYIAGNQBIGPSYEEFLKEKEKQKLKCLPPD 294
DB 266 -----PG-----RRGATPPWMIQDEEYIAGNQBIGPSYEEFLKEKEKQKLKCLPPD 311
QY 295 RVGANFDHSSRTSAGWLPSPFGRLXWTPLAVOTSTPKLKLQQRSSHIQKANHALYQL 354
DB 312 RVGANFDHSSRTSAGWLPSPFG---RVWN-----NGRRWQSRAPERGEVEVPWSEEQT 359
QY 355 PXGKPKSTK-PLLYLPPKPFILVFLRKQTVS 385
DB 360 PDPHTSRSLRIPAEVAPRSKRRCVLERKQPYS 391

RESULT 14
PCT-US04-12047-1081
; Sequence 1081, Application PC/TUS0412047
; GENERAL INFORMATION:
; APPLICANT: FIVE PRIME THERAPEUTICS, INC.
; TITLE OF INVENTION: HUMAN POLYPEPTIDES ENCODED BY POLYNUCLEOTIDES AND METHODS OF
; TITLE OF INVENTION: THEIR USE
; FILE REFERENCE: 08940.0021-00304
; CURRENT APPLICATION NUMBER: PCT/US04/12047
; CURRENT FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 1464
; PRIOR APPLICATION NUMBER: US 60/463,732
; PRIOR FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: US 60/463,708
; PRIOR FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: US 60/467,230
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/467,199
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/493,573
; PRIOR FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: US 60/493,577
; PRIOR FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: US 60/486,480
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/486,446
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/471,306
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/471,336
; PRIOR FILING DATE: 2003-05-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1081
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-12047-1081

Query Match 23.6%; Score 517; DB 1; Length 189;
Best Local Similarity 64.1%; Pred. No. 2.1e-37;
Matches 107; Conservative 18; Mismatches 36; Indels 6; Gaps 3;

```

```

QY 133 FLVTPQDYARFKKSMVKGGLDSYEEDKVKIKEMAAQIREVEQSRVRSVLE---PQAV 189
DB 24 YLASP-DYARFKKSLVEDLGSYEEDKVKIKEMAAQICEVEQSWQVQSVLEVGFPRI 82

```

Qy 190 PDPEGSSAPRSWKGMNSQVASSLQOPSNLDLPP--APELDWMTGPSLTFIGHQDIPGV 247
Db 83 QTPSASAGHSFSPWLEKDEPGSQFTAGLQGPATAPELDWMTGLSPIFIGHQDIPGV 142
Qy 248 GNIHSGATPPWMIQDEEYIAGNQEIGPSYEBFLKEKEKOKLKLPPD 294
Db 143 GNIHSGATPPWMIQDEECISGNQQTGPSYKEFIKEKEKELKELSPN 189

RESULT 15

PCT-US01-08631-44039
; Sequence 44039, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hysed, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 44039
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-08631-44039

Query Match 23.0%; Score 503; DB 1; Length 312;
Best Local Similarity 49.3%; Pred. No. 8.2e-36;
Matches 113; Conservative 14; Mismatches 38; Indels 64; Gaps 6;
Qy 40 TORQLKEAFERL-----LPQVEAARKAIRAAQVERYPVEHERCCWCLCCGCEVREHL 91
Db 108 TOPRLPAADWMPAATAITGTDVEAARKAIRAAQVERYPVEHERCCWCLCCGCEVREHL 167
Qy 92 SHGNLTVLYGGLLEHLASPEHKKATNKFENKAEVQMKKFLVTPQDYARF----- 143
Db 168 SHGNLTVLYGGLLEHLASPEHKKATNKFENKAEVQMKKFLVTPQDYARSHLYGTPMG 227
Qy 144 ----KKSVMKGL-----DSYEKEDKVIKEMAAQIREVQSQROEVVRSV-LBPOA 188
Db 228 SAVFSQACLEALLIIVPMGACGVSQEEES-----PARGSKDEPGEQVELKEEA 276
Qy 189 VPDPEGSSAPRSWKGMNSQVASSIQQPSNLDLPPAPELDWMTGPSLT 237
Db 277 EAPVEDGSQPP-----PPEPKGDATPGEKAT 303

Search completed: June 7, 2005, 12:19:34
Job time : 188.04 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 12:02:31 ; Search time 52.7607 Seconds
(without alignments)
2846.907 Million cell updates/sec

Title: US-09-155-676B-5

Perfect score: 2190

Sequence: 1 IGTVRRRSRIVDPAAXNS.....RGLESSISPTSRGGRTHF 417

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599520 seqs, 360203123 residues

Total number of hits satisfying chosen parameters: 1599520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/FCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	795	36.3	1115	9	US-09-915-543-17
2	795	36.3	1115	14	US-10-322-579-17
3	795	36.3	1115	15	US-10-664-859-17
4	198	9.0	104	10	US-09-746-783-206
5	148	6.8	389	15	US-10-424-599-209186
6	117.5	5.4	1040	16	US-10-425-115-306861
7	116	5.3	367	15	US-10-425-114-41619
8	116	5.3	363	15	US-10-425-114-69227
9	116	5.3	653	15	US-10-425-114-51821
10	112	5.1	160	16	US-10-425-115-219898
11	112	5.1	253	15	US-10-425-114-62537
12	112	5.1	253	15	US-10-425-114-67550

13	110.5	5.0	1878	15	US-10-607-631-20	Sequence 20, Appl
14	109.5	5.0	1109	16	US-10-437-963-133129	Sequence 133129,
15	109	5.0	4952	15	US-10-051-874-56	Sequence 56, Appl
16	109	5.0	5008	15	US-10-051-874-166	Sequence 166, App
17	109	5.0	5159	15	US-10-085-198-112	Sequence 112, App
18	109	5.0	5262	15	US-10-051-874-165	Sequence 165, App
19	107	4.9	447	16	US-10-425-115-368903	Sequence 368903,
20	106	4.8	270	14	US-10-105-021-2	Sequence 2, Appli
21	105.5	4.8	1879	15	US-10-607-631-8	Sequence 8, Appli
22	105	4.8	583	16	US-10-425-115-294295	Sequence 294295,
23	105	4.8	620	15	US-10-425-114-62454	Sequence 62454, A
24	105	4.8	2701	14	US-10-171-311-83	Sequence 83, Appl
25	104.5	4.8	526	15	US-10-094-749-1861	Sequence 1861, Ap
26	104.5	4.8	1094	14	US-10-032-585-7554	Sequence 7554, Ap
27	104.5	4.8	2093	16	US-10-437-963-128540	Sequence 128540,
28	104	4.7	1525	9	US-09-839-479-68	Sequence 68, Appl
29	104	4.7	1525	9	US-10-376-537-69	Sequence 69, Appl
30	104	4.7	1525	15	US-10-702-148-68	Sequence 68, Appl
31	104	4.7	5262	15	US-10-051-874-167	Sequence 167, App
32	103.5	4.7	753	16	US-10-755-889-358	Sequence 358, App
33	103.5	4.7	3224	10	US-09-315-355-34	Sequence 34, Appl
34	103	4.7	393	9	US-09-925-297-825	Sequence 825, App
35	103	4.7	905	13	US-10-114-893-127	Sequence 127, App
36	103	4.7	1056	15	US-10-275-595A-30	Sequence 30, Appl
37	103	4.7	1527	9	US-09-839-479-27	Sequence 27, Appl
38	103	4.7	1527	15	US-10-376-537-27	Sequence 27, Appl
39	103	4.7	1527	15	US-10-702-148-27	Sequence 27, Appl
40	103	4.7	1531	9	US-09-839-479-29	Sequence 29, Appl
41	103	4.7	1531	15	US-10-376-537-29	Sequence 29, Appl
42	103	4.7	1531	15	US-10-702-148-29	Sequence 29, Appl
43	102.5	4.7	563	16	US-10-425-115-286927	Sequence 286927,
44	102.5	4.7	964	16	US-10-437-963-181208	Sequence 181208,
45	102.5	4.7	1009	8	US-08-987-689A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-915-543-17
; Sequence 17, Application US/09915543
; Publication No. US20020086986A1
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich
; APPLICANT: FROESCH, Barbara
; APPLICANT: KRAMPS, Thomas
; APPLICANT: PETER, Oliver
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY
; FILE REFERENCE: Q60361
; CURRENT APPLICATION NUMBER: US/09/915,543
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,502
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Human lgs-1
US-09-915-543-17

Query Match	36.3%	Score 795;	DB 9;	Length 1115;
Best Local Similarity	55.6%;	Pred. NO. 3.5e-58;		
Matches 180;	Conservative 17;	Mismatches 67;	Indels 60;	Gaps 8
Qy	28	FFCGRGTTAASTORQLKEAFERLLPQVEAARKAIRAAQVRYVPEHERCCWCLCCGCEV	87	
		:		
Db	828	FFCGRGHVSRKHQRLKEALERLLPQVEAARKAIRAAQVRYVPEHERCCWCLCCGCEV	887	
		:		
Qy	88	REHLSHGNLTVLYGGLLEHLASPEHKATKNFHWENKAQVOMKEKFLVTPQDYARFKKSM	147	

Db 888 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 947
Qy 148 VKGLDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLEPQAVPDPPEGGSSAPRSWKGMNS 207
Db 948 VKGLDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLE-----TGPPR----- 990
Qy 208 QVASSLOQPSNLDLPPAPELDMWETGPSLTFIGHQDIPGVGNIHSGATPPMWIODEEYIA 267
Db 991 -YALTVRSPAVLSR-----RTLKSGAPFPQ--TPEAHPQ 1021
Qy 268 GNOETGSPSYEEFLKEKEQKLIK-KLPDRVGFANFDHSSRTSAGWLPSPGPRLEXWTPPLAV 326
Db 1022 ARCLCAPRRGALKPEPPGRTKLGVPPTHTRKARPHAARTSP-----RPRCTROAP--- 1072
Qy 327 QTSTPKLKL---QOXRSSHIOKKA 347
Db 1073 -NKTOSLQLAGKARTALHLQTKA 1095
RESULT 2
US-10-322-579-17
; Sequence 17, Application US/10322579
; Publication No. US20030114413A1
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich
; APPLICANT: FROESCH, Barbara
; APPLICANT: KRAMPS, Thomas
; APPLICANT: PETER, Oliver
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
; FILE REFERENCE: Q60361
; CURRENT APPLICATION NUMBER: US/10/322,579
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US/09/915,543
; PRIOR FILING DATE: 2001-07-27
; PRIOR FILING DATE: 2001-07-27
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 17
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Human lgs-1
US-10-322-579-17
Query Match 36.3%; Score 795; DB 14; Length 1115;
Best Local Similarity 55.6%; Pred. No. 3.5e-58;
Matches 180; Conservative 17; Mismatches 67; Indels 60; Gaps 8;
Qy 28 FFCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87
Db 828 FFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 887
Qy 88 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 147
Db 888 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 947
Qy 148 VKGLDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLEPQAVPDPPEGGSSAPRSWKGMNS 207
Db 948 VKGLDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLE-----TGPPR----- 990
Qy 208 QVASSLOQPSNLDLPPAPELDMWETGPSLTFIGHQDIPGVGNIHSGATPPMWIODEEYIA 267
Db 991 -YALTVRSPAVLSR-----RTLKSGAPFPQ--TPEAHPQ 1021
Qy 268 GNOETGSPSYEEFLKEKEQKLIK-KLPDRVGFANFDHSSRTSAGWLPSPGPRLEXWTPPLAV 326
Db 1022 ARCLCAPRRGALKPEPPGRTKLGVPPTHTRKARPHAARTSP-----RPRCTROAP--- 1072
Qy 327 QTSTPKLKL---QOXRSSHIOKKA 347
Db 1073 -NKTOSLQLAGKARTALHLQTKA 1095

RESULT 3
US-10-664-859-17
; Sequence 17, Application US/10664859
; Publication No. US20040038901A1
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich
; APPLICANT: FROESCH, Barbara
; APPLICANT: KRAMPS, Thomas
; APPLICANT: PETER, Oliver
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
; FILE REFERENCE: Q60361
; CURRENT APPLICATION NUMBER: US/10/664,859
; CURRENT FILING DATE: 2003-09-22
; PRIOR APPLICATION NUMBER: US/09/915,543
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,502
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 17
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Human lgs-1
US-10-664-859-17
Query Match 36.3%; Score 795; DB 15; Length 1115;
Best Local Similarity 55.6%; Pred. No. 3.5e-58;
Matches 180; Conservative 17; Mismatches 67; Indels 60; Gaps 8;
Qy 28 FFCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87
Db 828 FFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 887
Qy 88 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 147
Db 888 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 947
Qy 148 VKGLDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLEPQAVPDPPEGGSSAPRSWKGMNS 207
Db 948 VKGLDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLE-----TGPPR----- 990
Qy 208 QVASSLOQPSNLDLPPAPELDMWETGPSLTFIGHQDIPGVGNIHSGATPPMWIODEEYIA 267
Db 991 -YALTVRSPAVLSR-----RTLKSGAPFPQ--TPEAHPQ 1021
Qy 268 GNOETGSPSYEEFLKEKEQKLIK-KLPDRVGFANFDHSSRTSAGWLPSPGPRLEXWTPPLAV 326
Db 1022 ARCLCAPRRGALKPEPPGRTKLGVPPTHTRKARPHAARTSP-----RPRCTROAP--- 1072
Qy 327 QTSTPKLKL---QOXRSSHIOKKA 347
Db 1073 -NKTOSLQLAGKARTALHLQTKA 1095
RESULT 4
US-09-746-783-206
; Sequence 206, Application US/09746783
; Publication No. US20030044935A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John M.
; LaVallie, Edward R.
; Racie, Lisa A.
; Treacy, Maurice
; Spaulding, Vikki
; Agostino, Michael J.
; Howes, Steven H.
; Fechtel, Kim
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

Qy 117 NKFWENKAEVOMKEKELVTPQDYARFKKSMVGLDSVEEKEDKVIKEMAAQIREVQSR 176
Db 378 ERL--EDDERSYRLEKF-----DRLDIFQYIRHLEKEEQRVKQD---QVRRQERKN 427
Qy 177 QEVRSVLEPQAVDPDEGS-SAPRSWKGMNSQVASSLQOQSNLDLPPAPELDNMTGPS 235
Db 428 RDGRFKMLEHVA-----DGLNATRWRDYCAQIKDS----- 460
Qy 236 LTFIGHQDIPGVNIGHGATPPWMIQDEEYIAGNQIGPSY-EBFLKEKEKQKLKLP 292
Db 461 -----QSLAVASNTSGSTKELFDDV-----IEELGKQYQEDKIQIKVEVKGKIP 507

RESULT 10
US-10-425-115-219898
; Sequence 219898, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425.115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 219898
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_132134C.1.pap
US-10-425-115-219898

Query Match 5.1%; Score 112; DB 16; Length 160;
Best Local Similarity 28.3%; Pred. No. 0.27;
Matches 34; Conservative 17; Mismatches 39; Indels 30; Gaps 5;
Qy 249 NIHSATPPPMIODEEYIAGNQIGPSYEEFLKEKEKQKLKLPDRVGANFDHSSR--- 305
Db 22 NVHTGAPPPWLKANE-----HDPKNLSKSCGLSSRGKLRKLPNPNRVGAWAERRRAEM 76
Qy 306 -----TSAGWLPSPGPRLEXWTPPLAVOTSTPKLKLOQXRS--IQKANHAL 351
Db 77 EMEKQGEIVPATSDSSWLPNFG---SVW-----QSGTRKESRKEFEKSHKIHDTKSDHDL 128

RESULT 11
US-10-425-114-62537
; Sequence 62537, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425.114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62537
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700349560_FLI.pap
US-10-425-114-62537

Query Match 5.1%; Score 112; DB 15; Length 253;
Best Local Similarity 28.3%; Pred. No. 0.51;
Matches 34; Conservative 17; Mismatches 39; Indels 30; Gaps 5;
Qy 249 NIHSATPPPMIODEEYIAGNQIGPSYEEFLKEKEKQKLKLPDRVGANFDHSSR--- 305
Db 115 NVHTGAPPPWLKANE-----HDPKNLSKSCGLSSRGKLRKLPNPNRVGAWAERRRAEM 169
Qy 306 -----TSAGWLPSPGPRLEXWTPPLAVOTSTPKLKLOQXRS--IQKANHAL 351
Db 170 EMEKQGEIVPATSDSSWLPNFG---SVW-----QSGTRKESRKEFEKSHKIHDTKSDHDL 221

RESULT 12
US-10-425-114-67550
; Sequence 67550, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425.114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 67550
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3060-078-C12_FLI.pap
US-10-425-114-67550

Query Match 5.1%; Score 112; DB 15; Length 253;
Best Local Similarity 28.3%; Pred. No. 0.51;
Matches 34; Conservative 17; Mismatches 39; Indels 30; Gaps 5;
Qy 249 NIHSATPPPMIODEEYIAGNQIGPSYEEFLKEKEKQKLKLPDRVGANFDHSSR--- 305
Db 115 NVHTGAPPPWLKANE-----HDPKNLSKSCGLSSRGKLRKLPNPNRVGAWAERRRAEM 169
Qy 306 -----TSAGWLPSPGPRLEXWTPPLAVOTSTPKLKLOQXRS--IQKANHAL 351
Db 170 EMEKQGEIVPATSDSSWLPNFG---SVW-----QSGTRKESRKEFEKSHKIHDTKSDHDL 221

RESULT 13
US-10-607-631-20
; Sequence 20, Application US/10607631
; Publication No. US20040091901A1
; GENERAL INFORMATION:
; APPLICANT: Minion, F. Chris
; APPLICANT: Djordjevic, Steven P.
; TITLE OF INVENTION: Immunogenic Mycoplasma Hyopneumoninae
; FILE REFERENCE: 08411/035001
; CURRENT APPLICATION NUMBER: US/10/607.631
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 60/392,632
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1878
; TYPE: PRT
; ORGANISM: Mycoplasma hyopneumoninae
US-10-607-631-20

Query Match 5.0%; Score 110.5; DB 15; Length 1878;
Best Local Similarity 23.7%; Pred. No. 11;
Matches 71; Conservative 45; Mismatches 132; Indels 51; Gaps 14;

QY 92 SHGNLTLYVYGLLEHL-----ASPEHKATNKFWEWENKARVQMKFPL--VTQDYA 141
DB 414 SLGYNFLFDLASHLDYTLVSKAKIKQSSITKKLFIPLPIKISLSSILGDOENIKT 473

QY 142 RFKSKVMVGLDSEYEKE-DKVIKEMAAQI-REVEQSRQEVVRSVLEPQAVPDEGS-SA 198
DB 474 LFEKEVTFKLDNFRDVEIEKAFGLLYPGVNEELEQARREQASLEKEKAKGLKEFSQOK 533

QY 199 PRSWKGMNSQVASSLOQPSNLD--LPPAPELDWMTGPSITFGHODIPGVGNIHSGATP 256
DB 534 DENLKAINQ--DGLREDDNITELPENSPIQ-----QOEKAGLG---SSDPK 577

QY 257 PWMIQDEE----YIAGNQEIGPSYEFLEKEKQKLKLPDRVGFANFDHSSRTSAGMLP 312
DB 578 PYMIKDVQNRYYLAKSQ-----IQELIKAKDYTKLAKLSNRHTYNSLRLEQ---LF 629

QY 313 SFGPRLEXTWPLA---VQTSTPKLKLQQRSS-----HIQKXANHALYQLPXGXKPKST 363
DB 630 EVNPRIPSSRDIEAKFVLDKTEKNKYQIYSSASPAFQNKWSLFGYYRYLLGLDLPKQT 688

RESULT 14
US-10-437-963-133129
; Sequence 133129, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 133129
; LENGTH: 1109
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_3502C.1.pap
US-10-437-963-133129

Query Match 5.0%; Score 109.5; DB 16; Length 1109;
Best Local Similarity 22.2%; Pred. No. 6.3;
Matches 76; Conservative 43; Mismatches 126; Indels 97; Gaps 18;

QY 90 HLSHGNLTLYVYGLLEHLASPEHK-KATNKFWEWEN-----KAEVQMKFEL----- 134
DB 8 HAEKKLNLLQCNQDQEKVAFATHQLQGASAWDNHMATRPDDVRTDAEKQKFMAGLD 67

QY 135 --VTQ-----DYARPKSMVGLDSEYEKEKDVIKEMAAQIREVEQSRQEVVRSVLEPQA 188
DB 68 DELTNQLTSGDYADFERLVDAIROEDQNRKMDRKRKAAQFRAHQGSHQ-----RPRF 120

QY 189 VPDPEEGSSA--PRSWKGMNSQVASSLOQPSNLDLPPAPELDWMTGPSITFGHODIPG 246
DB 121 TPGQGGGTMTLVIRYRFN-----PSN-----FHQAS-----GSQDQH 156

QY 247 VGNTHSGATPMMIODEBYIAGNQIGPSYEFLEKEKQKLKLPDRVGFANFDHSSRT 306
DB 157 -GQLNRGAAAP-----RPPVAPAQSGPS-----AQAKKETGKPGSCFNCG--- 195

QY 307 SAGWLPSTGPRLEXTWPLAVQTSTPKLKLQQRSSHIQKXANHALYQLPXGXKPKSTKPL 366
DB 196 ELGHNLADKCPKPRRAGP-----RFVQARVNHASAEAEAAPEVVLTGTFPVNSIPA 245

QY 367 LYLPPKFFIIVFLRQQTYSFXFNKVLFPFGLRGL-ESSLSIP 407
DB 246 -----TVLFDSGATHSFISKK---FVGMHGLVREELSTP 276

RESULT 15
US-10-051-874-56
; Sequence 56, Application US/10051874
; Publication No. US2004000557A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Saeha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Macdougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glennnda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: USING THE SAME
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051,874
; CURRENT FILING DATE: 2002-09-25
; PRIOR FILING DATE: 2002-09-25
; PRIOR FILING DATE: 2001-02-14
; PRIOR FILING DATE: 2001-02-14
; PRIOR FILING DATE: 2001-09-27
; PRIOR FILING DATE: 2001-09-27
; PRIOR FILING DATE: 2001-01-18
; PRIOR FILING DATE: 2001-01-18
; PRIOR FILING DATE: 2001-02-28
; PRIOR FILING DATE: 2001-02-28
; PRIOR FILING DATE: 2001-01-18
; PRIOR FILING DATE: 2001-01-18
; PRIOR FILING DATE: 2001-03-16
; PRIOR FILING DATE: 2001-03-16
; PRIOR FILING DATE: 2001-05-17
; PRIOR FILING DATE: 2001-05-17
; PRIOR FILING DATE: 2001-10-18
; PRIOR FILING DATE: 2001-10-18
; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/261,376
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 4952
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-051-874-56

Query Match 5.0%; Score 109; DB 15; Length 4952;
Best Local Similarity 20.9%; Pred. No. 54;
Matches 80; Conservative 51; Mismatches 149; Indels 102; Gaps 15;

Qy	101	GGLLEHLASPE-----HKKATNKFWMENKA-----EVQMKKEFLVTPQDYAR	142
Db	2609	GGPPAHLITPSPLSGPGSSGLEKFELESALTLPGGPAASGDDELDMESSLVASE----	2664
Qy	143	FKSMVKGLDSYEKEDKVIKEMAAQIREVEQSREQEVVRSVLEP-----QAVDPPEGSS	197
Db	2665	-LPLLIEDLLEHEKELQKQQLSAQLQPAQOQOQOQOQHSLLPAPQAPQAMSLPHEGSS	2723
Qy	198	APRSWKGMNSOVA-----SSLQPSNLDLPAPELDMETGPSLTFIGHQ-----	242
Db	2724	P--SLAGSQQLSLGLAVARQPLPQPLMPTQPPAHALQ--ORLAPSMWVSNQGHMLSGQ	2780
Qy	243	-----DIPGVNIHSGATPPMMIQDEEYIAGNQIGPSY-----EEF	279
Db	2781	HGGAGLVPOQSSQPVLSQKPMGTMPSPCMKPPQLAMQQLANSFPFDTDLDKFAAEDI	2840
Qy	280	LKEKEKQK-----LKKLPDRVGNANFDHSSRTSAGWLPSFGPRLEXWTPPLAVQTSTPKL	333
Db	2841	IGPIAKAMVALKGIKKV-----MAQGSIGVAPGMNRQ-----QVSLLAQ	2880
Qy	334	KLQQRSSHIOKXANHALYQLPGXKPKSTKPLLVLPKPF--IIVFLRKOTYSFIKFN-	390
Db	2881	RLSGPSSDLQNHVAAGSGQERSAGDFSQPRP---NPPTFAQGVINEADQRYEEWLFHT	2937
Qy	391	-KVLFFGLRGLSLSIPSTSR	411
Db	2938	QQLLMQLKVLREQIGVHRKSR	2959

Search completed: June 7, 2005, 12:25:33
Job time : 54.7607 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query Match	Length	DB	ID	Description
1	108.5	5.0	1225	4	US-09-501-171-4	Sequence 4, Appli
2	108.5	5.0	1225	4	US-09-949-016-6063	Sequence 6063, Ap
3	107.5	4.9	1540	4	US-09-949-016-7037	Sequence 7037, Ap
4	104	4.7	1525	4	US-09-418-710-69	Sequence 69, Appl
5	104	4.7	1525	4	US-09-839-479-68	Sequence 68, Appl
6	103.5	4.7	578	3	US-09-066-046-6	Sequence 6, Appli
7	103.5	4.7	753	4	US-09-949-016-6676	Sequence 6676, Ap
8	103.5	4.7	3224	4	US-08-705-660-34	Sequence 34, Appl
9	103.5	4.7	3224	3	US-08-989-045-34	Sequence 34, Appl
10	103.5	4.7	3224	4	US-09-538-092-1161	Sequence 1161, Ap
11	103.5	4.7	3224	4	US-09-315-355A-33	Sequence 34, Appl
12	103	4.7	1527	4	US-09-418-710-27	Sequence 27, Appl
13	103	4.7	1527	4	US-09-839-479-27	Sequence 27, Appl
14	103	4.7	1531	4	US-09-418-710-29	Sequence 29, Appl
15	103	4.7	1531	4	US-09-839-479-29	Sequence 29, Appl
16	102.5	4.7	667	4	US-09-949-016-7759	Sequence 7759, Ap
17	102.5	4.7	777	4	US-09-949-016-8706	Sequence 8706, Ap
18	102.5	4.7	777	4	US-09-949-016-8707	Sequence 8707, Ap
19	102.5	4.7	1009	2	US-08-357-642A-1	Sequence 1, Appli
20	102.5	4.7	1009	2	US-08-460-626-1	Sequence 1, Appli
21	102.5	4.7	1014	4	US-09-949-016-11533	Sequence 11533, A
22	101	4.6	2101	1	US-08-466-390-4	Sequence 4, Appli
23	101	4.6	2101	1	US-08-470-950-4	Sequence 4, Appli
24	101	4.6	2101	1	US-08-467-781-4	Sequence 4, Appli
25	101	4.6	2101	1	US-08-195-487-4	Sequence 4, Appli
26	101	4.6	2101	2	US-08-483-924-4	Sequence 4, Appli
27	101	4.6	2101	3	US-09-452-294-1	Sequence 1, Appli

[illegible]

QY 25 EGPFCCGRTTAATQOLKEAPERLLPQVEAARKAIRAAQVERVYVPEHRCWCWCLCCG 84
D 19 BEPF-----FTIPHTQ-----EAF-RTREEYE-----ARLERY-----SERIWTCKSTG 56
QY 85 CEVREHLHGHLNLTLYGGLLEHLASPEHKKATNKFWMENKAEVO--MKEKFLVTPQDYAR 142
D 57 -----SSQLT-----HKEA-----WEEQEVALLKEEF---PANYEK 86
QY 143 FKSMV-----KGLDS-----YEEKDKVIKEMAAQIREVEQSRQE 178
D 87 LVLEWVHHTASLEKLVDTAWLEIMTKYAVGEECDFEVGEKMKLVKIVKIHPLEK----- 215
QY 179 VVRSVLPEQAVDPDEEGS-SAPRSWKGMNSQVASSLQO-----PAPELDMWETGPSLT---FI 239
D 143 -----VDEEATEKSGDACSPPSSDKENSSQIAQHQKKTETVVKKEDEGRRESINDRARRS 197
QY 216 -----PSNLD-----LP-----PAPELDMWETGPSLT---FI 239
D 198 PRKLPSTLKKGERKWAPPKFLPHKYDVKLQNEDKLIISNVPADSLIRTERPKNKEIVRYFI 257
QY 240 GHODIPGVGNHSGATPPMMIQDEEYIAGNOEIGPSYEEFLKEKEK-----QKLKKL 291
D 258 RHNA-----RAGTENAPWVVEDE--LVKKYSLPSKFSDFLLDPYKMTLNPSTKRKNTG 311
QY 292 PPDR---VGANFDSHRTSAGWLPSPGPRLEW-----TPLAVQTS-----TPKL 333
D 312 SPDRKPSKSKTDNSLSL-----PLNPKL--WCHVHLKKSLSGSLPKVKNSKNSKSPDEE 364
QY 334 KLOQ-----XRSISHOKKANHALYQLPXGKXP--KSTKPL 366
D 365 HLEEMMKQMSPNKLTNPHIPKKG-----PPAKKPGKHSKPL 402

RESULT 5

US-09-839-479-68
; Sequence 68, Application US/09839479
; Patent No. 6727222
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-042002
; CURRENT APPLICATION NUMBER: US/09/839,479
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 09/418,710
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01783
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JP 9/310027
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: JP 9/116570
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 1525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-839-479-68

Query Match 4.7%; Score 104; DB 4; Length 1525;

Best Local Similarity 21.3%; Pred. No. 1.5;

Matches 99; Conservative 53; Mismatches 109; Indels 204; Gaps 29;

QY 25 EGPFCCGRTTAATQOLKEAPERLLPQVEAARKAIRAAQVERVYVPEHRCWCWCLCCG 84
D 19 BEPF-----FTIPHTQ-----EAF-RTREEYE-----ARLERY-----SERIWTCKSTG 56
QY 85 CEVREHLHGHLNLTLYGGLLEHLASPEHKKATNKFWMENKAEVO--MKEKFLVTPQDYAR 142
D 57 -----SSQLT-----HKEA-----WEEQEVALLKEEF---PANYEK 86
QY 143 FKSMV-----KGLDS-----YEEKDKVIKEMAAQIREVEQSRQE 178

Db 87 LVLEWVHHTASLEKLVDTAWLEIMTKYAVGEECDFEVGEKMKLVKIVKIHPLEK----- 142
QY 179 VVRSVLPEQAVDPDEEGS-SAPRSWKGMNSQVASSLQO----- 215
D 143 -----VDEEATEKSGDACSPPSSDKENSSQIAQHQKKTETVVKKEDEGRRESINDRARRS 197
QY 216 -----PSNLD-----LP-----PAPELDMWETGPSLT---FI 239
D 198 PRKLPSTLKKGERKWAPPKFLPHKYDVKLQNEDKLIISNVPADSLIRTERPKNKEIVRYFI 257
QY 240 GHODIPGVGNHSGATPPMMIQDEEYIAGNOEIGPSYEEFLKEKEK-----QKLKKL 291
D 258 RHNA-----RAGTENAPWVVEDE--LVKKYSLPSKFSDFLLDPYKMTLNPSTKRKNTG 311
QY 292 PPDR---VGANFDSHRTSAGWLPSPGPRLEW-----TPLAVQTS-----TPKL 333
D 312 SPDRKPSKSKTDNSLSL-----PLNPKL--WCHVHLKKSLSGSLPKVKNSKNSKSPDEE 364
QY 334 KLOQ-----XRSISHOKKANHALYQLPXGKXP--KSTKPL 366
D 365 HLEEMMKQMSPNKLTNPHIPKKG-----PPAKKPGKHSKPL 402

RESULT 6

US-09-066-046-6
; Sequence 6, Application US/09066046A
; Patent No. 6204252
; GENERAL INFORMATION:
; APPLICANT: MURPHY, Cheryl
; STOREY, James
; BELTZ, Gerald A.
; COUGHLIN, Richard T.
; TITLE OF INVENTION: CHARACTERIZATION OF GRANULOCYTIC
; EHRlichia AND METHODS OF USE

NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESS: HALE AND DORR LLP
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: United States
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/066,046A
FILING DATE: 24-Apr-1998
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Superko, Colleen
REGISTRATION NUMBER: 39,850
REFERENCE/DOCKET NUMBER: 106,941,155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-5000
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 578 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-066-046-6

Query Match 4.7%; Score 103.5; DB 3; Length 578;

Best Local Similarity 20.8%; Pred. No. 0.4;

Matches	64;	Conservative	51;	Mismatches	114;	Indels	79;	Gaps	13;
QY	13	DPRAAXNSIRAHEGFPFCGRGTTAASTORQKAEFERLLPOVEAARAKAIRAAQVERVY-	71						
Db	10	DYRIGSEYISSGDDGYEGCADCDASTNAYSID-----KCRVVGRTWRPSELVLYVG	62						
QY	72	PEHERCCWCLCGCCEVREHLHSGNLT-----VLYGGLLEHLASPEHK	113						
Db	63	DEHVAC-----RDVASGMHGNLPGKVYFIEAEGRAATAEGGVVTTVVEALSIVQEE	115						
QY	114	KATNKFWNENKAEVOMKEKFLV-----TPQ--DYARFKSMVKGLDSYEBKEDEK	160						
Db	116	ECTGYML-----INAPKEAVRRFFKIEKSAABEPQTVDPSPVSESATGSGVDTQEEQE--	167						
QY	161	VTKENAAQIREVQESQRQVRSVLQPCQAVDPPEGSG-----APRSWKGNNS--QVASSL	213						
Db	168	IDQAPALIEEVETEEQEV-----ILEBGTLDLEQFVAPVPAEALPGVEAAEAIVPSL	223						
QY	214	QQPSNLDLPPALDMMETGPSLTF-----IGHQDIPGVGNIHSGATPPWMIQDE	263						
Db	224	EENKLOEVVVAPAEQAQLESAPESVAPAPQESTVLGVAEGDLKSEVSVEANADVP-----QK	279						
QY	264	EYIAGNQE	271						
Db	280	EVISGCOE	287						

```

RESULT 7
US-09-949-016-6676
; Sequence 6676, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6676
; LENGTH: 753
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6676

```

Query Match	4.7%	Score 103.5;	DB 4;	Length 753;
Best Local Similarity	22.0%;	Pred. No. 0.59;		
Matches	84;	Conservative 48;	Mismatches 149;	Indels 101;
Gaps	19;			
19 NSIRAH-----EGPFCG-RGTFTAASTORLK-----EAFERLLPQVE	56			
195 HSLRIHVRVHTKEPFCDVQGEKAFNTYLRKAHQRLHTKTFNCESEGCKYFTTLS	244			
57 AARKAIRAAQVER-----YVPEH--BRCWCWLCGCE-----	86			
245 DLRKHIRTHTGKPPFRCDHDCGKAFAAASHHLKTHVTRTGTGPRFFCPGNGCEKTFSTQY	304			
87 -VREHL-SHGNTLVLYGGLLEHLASPBHK-----ATNKFWEKASVOMKEKFL	134			
305 SLKSHMKCHDNKGHSYNALPQNGSEDTNHSCLCLDLSTLSDSELNENSTTQQDILST	364			
135 VTPODYARFKSMVKGLDSTVEEKEDKVIKEMAAQIREVCGSQROEVVRSVLEPQAVDPDEE	194			
365 ISP---AIIPESM-----FQNSDDTAIQBDPQQTASLTSEFNGDAESVD---YP-EST	411			
195 GSSAPRS-----WKGMSQVASSIQPNSLDLPAPPELDWMTGTSLTPIGHODIPGVGN	249			

```

Db      412 GNSASLSLPLVLQFGLSEPPQPLLPAAGAPSPAPPSL-----GP-----GSQQ-AAFGN 460
Qy      250 IHSGATPPWMIODEY-IAGNQELGPSYEELKEKEKOKLKLPPDRVGANFDHSSRTSA 308
Db      461 -----PPALLQPPFVFPVPHSTQFAANHQEFPLPQAPQ-PIVPGLSVVAGASASAAAAVA 513
Qy      309 GWLPSFGPRLEXWTPLAVQTST 330
Db      514 SAVAAPPPQSTTEPLPAMVQT 535

RESULT 8
US-08-705-660-34
; Sequence 34, Application US/08705660
; Patent No. 5858683
; GENERAL INFORMATION:
; APPLICANT: KEESEE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-JYE
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: DETECTION OF CERVICAL CANCER
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,660
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: GREENHALGH, DUNCAN A
; REGISTRATION NUMBER: 38,678
; REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3224 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-705-660-34

```

Query Match	4.7%	Score 103.5	DB 2	Length 3224
Best Local Similarity	18.2%	Pred. No. 5.3		
Matches	74	Conservative 69	Mismatches 132	Indels 131
Gaps	17			
QY	48	FERLLPOVEAARKA-----	-----IRAAOVERVYPHERCCWCLCCGCEVRE	89
Db	606	WKKVLPLLKTIKKNSIPEDLPKFHFSVDIQASEIVY-----	-----BEDA	650
QY	90	HLSHGNLTVLYGLLEHSLASPEHKATKNFW-----	-----WENKA-----EVOMKEKF	133
Db	651	HITFAILDVANGNIEDAVTAFEIKSVVSNLIALIPHRKAEDLNDALSPSEOECKNY		710
QY	134	LVTPODYARFKSMVKGLDSEYEDKVKIKEMAAQIREVEQSRQSVRVSVLEPQAVDPPE		193
Db	711	LRKTRDY-----LIKTIID-SDSNLSVVKKLPVLESV-----KEMLSVM--	-----CELEDYS	757
QY	194	EGSSAPR--SWKGMNSOAVSSLQOPSNLDLPAPPELDMWETGTPSLTFIGHODIPGVGNIH		251

Db 758 EGGPLYKNGSLRNADSEIKRSTPSPTRYSLSPSKSYK-----Y 795
Qy 252 SGATPPMMIODEEYIAGNQEIGPSYEFLEKEKQKL-----KKLPPDRVGA 298
Db 796 SPKTPPRAEDQNSLL--KMICQVEAIKKEMQELKLNSSNSASPHRWPTENYGPDSVPD 853
Qy 299 NFDHSSSTSAGWLPSPGPRLEXWTPLAVQSTPKLKLQQRSSHIQKANHLYQLPKGX 358
Db 854 GY-QGSQTFHG-----APLTVAATGPGSV-----YYSQSPAYNSQYLLRPAA 893
Qy 359 KPKSTKPLLY---LPPKFFIIVFLRK-----QTSYFIKFNKVL 394
Db 894 NVPTTKGPVYGMNRLPPQOHYIAYPQOMHTPPVQSSACMFQSEMY 939
RESULT 9
US-08-989-045-34
; Sequence 34, Application US/08989045
; Patent No. 6027905
; GENERAL INFORMATION:
; APPLICANT: KEESEE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-JYE
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; DETECTION OF CERVICAL CANCER
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,045
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GREENHALGH, DUNCAN A
; REGISTRATION NUMBER: 38,678
; REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3224 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-989-045-34
Query Match 4.7%; Score 103.5; DB 3; Length 3224;
Best Local Similarity 18.2%; Pred. No. 5.3;
Matches 74; Conservative 69; Mismatches 132; Indels 131; Gaps 17;
Qy 48 FERLLPQVEAARKA-----IRAAQVERVYVPEHERCCWCLCGCEVRE 89
Db 606 WKVPLPLKIIKKNSIPEPIDPLFKHFHSDVIOASEIVEY-----BEDA 650
Qy 90 HLSHGNLTVLYGGLLEHLASPEHKKATNKW-----WENKA-----EVQMEKF 133
Db 651 HITAILDAVNGNIEDAVTAFESIKSVSVYWNLLALIFHRKAEDIENDALSPEEQECKNY 710
Qy 134 LVTPQDYARFKKSMVKGGLDSYEEKEDKVIKEMAAQIREVEQSROEVRSVLEPQAVDPE 193
Db 711 LRKTRDY-----LIKIIDD--SDSNLSVVKLPVPLESV-----KEMLSNVM--QLEEDYS 757
US-08-989-045-34

Qy 194 EGSSAPR--SWKGNMSQVASSLQOPSNLDLPPAPELDWMETGPSLTFIGHODIPGVGNIH 251
Db 758 EGGPLYKNGSLRNADSEIKRSTPSPTRYSLSPSKSYK-----Y 795
Qy 252 SGATPPMMIODEEYIAGNQEIGPSYEFLEKEKQKL-----KKLPPDRVGA 298
Db 796 SPKTPPRAEDQNSLL--KMICQVEAIKKEMQELKLNSSNSASPHRWPTENYGPDSVPD 853
Qy 299 NFDHSSSTSAGWLPSPGPRLEXWTPLAVQSTPKLKLQQRSSHIQKANHLYQLPKGX 358
Db 854 GY-QGSQTFHG-----APLTVAATGPGSV-----YYSQSPAYNSQYLLRPAA 893
Qy 359 KPKSTKPLLY---LPPKFFIIVFLRK-----QTSYFIKFNKVL 394
Db 894 NVPTTKGPVYGMNRLPPQOHYIAYPQOMHTPPVQSSACMFQSEMY 939
RESULT 10
US-09-538-092-1161
; Sequence 1161, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormatter Version 0.9
; SEQ ID NO 1161
; LENGTH: 3224
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P49792
US-09-538-092-1161

Query Match 4.7%; Score 103.5; DB 4; Length 3224;
Best Local Similarity 18.2%; Pred. No. 5.3;
Matches 74; Conservative 69; Mismatches 132; Indels 131; Gaps 17;
Qy 48 FERLLPQVEAARKA-----IRAAQVERVYVPEHERCCWCLCGCEVRE 89
Db 606 WKVPLPLKIIKKNSIPEPIDPLFKHFHSDVIOASEIVEY-----BEDA 650
Qy 90 HLSHGNLTVLYGGLLEHLASPEHKKATNKW-----WENKA-----EVQMEKF 133
Db 651 HITAILDAVNGNIEDAVTAFESIKSVSVYWNLLALIFHRKAEDIENDALSPEEQECKNY 710
Qy 134 LVTPQDYARFKKSMVKGGLDSYEEKEDKVIKEMAAQIREVEQSROEVRSVLEPQAVDPE 193
Db 711 LRKTRDY-----LIKIIDD--SDSNLSVVKLPVPLESV-----KEMLSNVM--QLEEDYS 757
Qy 194 EGSSAPR--SWKGNMSQVASSLQOPSNLDLPPAPELDWMETGPSLTFIGHODIPGVGNIH 251
Db 758 EGGPLYKNGSLRNADSEIKRSTPSPTRYSLSPSKSYK-----Y 795
Qy 252 SGATPPMMIODEEYIAGNQEIGPSYEFLEKEKQKL-----KKLPPDRVGA 298
Db 796 SPKTPPRAEDQNSLL--KMICQVEAIKKEMQELKLNSSNSASPHRWPTENYGPDSVPD 853
Qy 299 NFDHSSSTSAGWLPSPGPRLEXWTPLAVQSTPKLKLQQRSSHIQKANHLYQLPKGX 358
Db 854 GY-QGSQTFHG-----APLTVAATGPGSV-----YYSQSPAYNSQYLLRPAA 893

QY 359 KPKSTKPLLY-----LPPKFFIIVFLRK-----QTVSFXFNKVL 394
Db 894 NVTPTKGPVGNRLPPQOHYAYPQOMHTPPVQSSACMFQSEMY 939

RESULT 11

US-09-315-355A-34
; Sequence 34, Application US/09315355A

; Patent No. 6803189

; GENERAL INFORMATION:

; APPLICANT: Keese, Susan

; APPLICANT: Obar, Robert

; APPLICANT: Wu, Ying-Yue

; TITLE OF INVENTION: Methods for the Detection of Cervical Cancer

; FILE REFERENCE: MTP-023DV2

; CURRENT APPLICATION NUMBER: US/09/315,355A

; CURRENT FILING DATE: 1999-05-17

; PRIOR APPLICATION NUMBER: US 08/989,045

; PRIOR FILING DATE: 1997-12-11

; PRIOR APPLICATION NUMBER: US 08/705,660

; PRIOR FILING DATE: 1996-08-30

; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 34

; LENGTH: 3224

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-315-355A-34

Query Match 4.7%; Score 103.5; DB 4; Length 3224;
Best Local Similarity 18.2%; Pred. No. 5.3;

Matches 74; Conservative 69; Mismatches 132; Indels 131; Gaps 17;

QY 48 BELLPOVEAARKA-----IRAAQVRYVPEHERCCWCLCCCEVRE 89
Db 606 WKVVLPLKLIKKNKNSIEPIDLPKFHSDVDIQASEIVEY-----BEDA 650

QY 90 HLHSHGLNLTLYGLLEHLASPEHKATNKFW-----WENKA-----EVQMKKF 133
Db 651 HITFAILDVANGNIEDAVTAFESIKSVVSNWNLALIFHRKAEDINDALSPEEQECKNY 710

QY 134 LVTPODYARFKSMVKGLDSYEKEDKVIKEMAAQIREVEQSRQVRSVLEPQAVDPPE 193
Db 711 LKTRDY-----LIKIIDD--SDSNLSVVVKLPVPLESV-----KEMLSVM--QLEDEYS 757

QY 194 EGSSAPR--SWKGMNSQVASSLQOPSNLDLPPAPELDNMTGSLTFIGHQDIPGVGNH 251
Db 758 EGGPLYKNGSLRNAUSEIKRSTPSTRYLSLSPSKYK-----Y 795

QY 252 SGATPPWMIQDEEYIAGNQEIGPSYEFLKEKEKOKL-----KKLPPDRVGA 298
Db 796 SPKTPPRWAEDQNSLL--KMIQQQVEAIKEMQELKLNSSNSASPHRWPTENYGPDSVPD 853

QY 299 NFDHSSRTSAGLWPSFGFRLEXTWTPAVQSTPKLQOXRSSHOKKANHALYOLPXGX 358
Db 854 GY-QGSQTFHG-----APLTVATTGPSV-----YYSQSPAYNSQVLLRPA 893

QY 359 KPKSTKPLLY-----LPPKFFIIVFLRK-----QTVSFXFNKVL 394
Db 894 NVTPTKGPVGNRLPPQOHYAYPQOMHTPPVQSSACMFQSEMY 939

RESULT 12

US-09-418-710-27
; Sequence 27, Application US/09418710

; Patent No. 6596482

; GENERAL INFORMATION:

; APPLICANT: Jones, Michael H.

; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR

; FILE REFERENCE: 06501-042001

; CURRENT APPLICATION NUMBER: US/09/418,710

; CURRENT FILING DATE: 1999-10-15

; PRIOR APPLICATION NUMBER: PCT/JP98/01783

; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JP 9/310027
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: JP 9/116570
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 1527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-418-710-27

Query Match 4.7%; Score 103; DB 4; Length 1527;

Best Local Similarity 21.3%; Pred. No. 1.9; Mismatches 53; Indels 204; Gaps 29;

QY 25 EGPFFCGRTTAASTQRLKEAFERLLPQVEAARKAIRAAQVRYVPEHERCCWCLCCG 84
Db 19 EEPF-----FTIPHTQ---EAF-RTREEYE-----ARLERY---SERIWTCKSTG 56

QY 85 CEVREHLSHGNLTLYGLLEHLASPEHKATNKFWENKAEVQ--MKEKFLVTQDYAR 142
Db 57 -----SSQLT-----HKEA-----WEEQEVAEILLKEEF---PAWYEK 86

QY 143 FKSMV-----KGLDS-----YBEKEDKVIKEMAAQIREVEQSRQE 178
Db 87 LVLENVHINTASLEKLVDTAWLEIMTKYAVGEECDFFVGKMKLVKIVKIHLEK---- 142

QY 179 VVRSVLEPQAVDPPEGS-SAPRSKGMNSQVASSLQO----- 215
Db 143 -----VDEATEKKSDGACDSSDKENSSQIAQDHQKKTWVKEDEGRRESINDRARS 197

QY 216 -----PSNLD-----LP-----PAPELDNMTGSLT-----FI 239
Db 198 PRKLFTSLKGERKWAPKFLPHKYDVKLQNEDKIISNVPADSLIRTPPNKEIVRYFI 257

QY 240 GHQDIPGVGNTHSGATPPWMIQDEEYIAGNQEIGPSYEFLKEKEK-----QKCLK 291
Db 258 RHNL-----RAGTENAPWVEDE--LVKKYSLPSKSDFLDDPYKMTLNPSTYKRNKG 311

QY 292 PPDR---VGANFDSHSSRTSAGLWPSFGFRLEKW-----TPLAVQTS-----TPKL 333
Db 312 SPDRKPSKSKTDNSSLSS-----PLNPKL--WCHVHLKSLSGSPLKVKNSKNSKSPDEE 364

QY 334 KLQ-----XRSISHIOKANHALYOLPXGKP--KSTKPL 366

Db 365 HLEEMKMSPNKLTNPHIPKKG-----PPAKKPGKSDKPL 402

RESULT 13

US-09-839-479-27
; Sequence 27, Application US/09839479

; Patent No. 6727222

; GENERAL INFORMATION:

; APPLICANT: Jones, Michael H.

; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR

; FILE REFERENCE: 06501-042002

; CURRENT APPLICATION NUMBER: US/09/839,479

; CURRENT FILING DATE: 2001-04-20

; PRIOR APPLICATION NUMBER: US 09/418,710

; PRIOR FILING DATE: 1999-10-15

; PRIOR APPLICATION NUMBER: PCT/JP98/01783

; PRIOR FILING DATE: 1998-04-17

; PRIOR APPLICATION NUMBER: JP 9/310027

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: JP 9/116570

; PRIOR FILING DATE: 1997-04-18

; NUMBER OF SEQ ID NOS: 72

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 27

; LENGTH: 1527

; TYPE: PRT

[illegible]

Mon Jun 13 13:27:15 2005

```
QY 216 -----PSNLD-----LP-----PAPELDWMTGPSLT-----FI 239
Db 198 PRKLPSTLKGGERKWAPPKFLPHKYDVKLQNEDKIISNVPADSLIRTERPPNKIEIVRYFI 257
QY 240 GHQDIPGVGNHSGATPPWMIQDEBEYIAGNQBIGPSVEEFLKEKEK-----QKLKKL 291
Db 258 RHNAL-----RAGTGENAPWVEDE--LVKKYSLPSKFSDFLLDPYKYMTLNPSTKRKNTG 311
QY 292 PPDR---VCANFDHSSRTSAGWLPSPGPRLEW-----TPLAVQTS-----TPKL 333
Db 312 SPDRKPSKSKTDNSLSS-----PLNPKL--WCHVHLKKSLSGSPKVKNSKNSKSPEE 364
QY 334 KLOO-----XRSSHIQKANHLYQLPXGXP--KSTKPL 366
Db 365 HLEEMMKMSPNKLHTNPHIPKG-----PPAKKPGKHSDKPL 402
```

Search completed: June 7, 2005, 12:04:51
Job time : 19.3155 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:18:24 ; Search time 13067.6 Seconds
(without alignments)
17042.190 Million cell updates/sec

Title: US-09-155-676B-6
Perfect score: 4596
Sequence: 1 agcggggggactgtgccgtg.....gatgacaaatgttaaaaaa 4596

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ets.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4596	100.0	4596	6	A66650 Sequence 6
2	4578.4	99.6	4596	6	AX429240 Sequence
3	4578.4	99.6	4596	6	AX774978 Sequence
4	4578.4	99.6	4596	9	HSNIK
5	4438.4	96.6	4584	6	CQ723104 Sequence
6	4249.4	92.5	4452	9	BC035576 Sequence
7	3149.6	68.5	3156	6	AR062290 Sequence
8	3149.6	68.5	3156	6	AR063065 Sequence
9	3149.6	68.5	3156	6	AR068515 Sequence
10	3149.6	68.5	3156	6	BD062401 Sequence
11	2837.6	61.7	2844	6	AR429673 Sequence
12	2673.4	58.2	3152	6	C0845924 Sequence
13	2673.4	58.2	3152	9	AK131438 Sequence
14	2054.6	44.7	2829	10	AF143094 Sequence
15	1851.2	40.3	2631	6	A66647 Sequence 3
16	1556.2	33.9	48032	9	AC003963 Sequence
17	1556.2	33.9	96875	9	AC008105 Sequence
18	1556.2	33.9	100836	2	AC087298 Sequence
19	1544.2	33.6	142326	2	AC024047 Sequence

c	20	620	13.5	71599	2	AC139008	AC139008 Homo sapi
c	21	504.4	11.0	162192	2	AC150796	AC150796 Bos tauru
c	22	480.2	10.4	223438	10	AL662804	AL662804 Mouse DNA
c	23	479	10.4	111803	2	AC136172	AC136172 Rattus no
c	24	477	10.4	239464	2	AC103079	AC103079 Rattus no
c	25	476	10.4	476	6	AX778201	AX778201 Sequence
c	26	459.6	10.2	70748	2	AC137575	AC137575 Homo sapi
c	27	459.6	10.2	142326	2	AC024047	AC024047 Homo sapi
c	28	440.2	9.6	698	11	BV166476	BV166476 MAP3K14.9
c	29	405.8	8.8	70748	2	AC137575	AC137575 Homo sapi
c	30	349.4	7.6	2475	6	AX834630	AX834630 Sequence
c	31	349.4	7.6	2475	9	AK097219	AK097219 Homo sapi
c	32	333	7.2	337	11	G27543	G27543 human SFS S
c	33	298.2	6.5	362	6	AX886720	AX886720 Sequence
c	34	298.2	6.5	362	6	BD026330	BD026330 Sequence
c	35	291.6	6.3	6673	6	AX458595	AX458595 Sequence
c	36	243.6	5.3	6673	6	AX458596	AX458596 Sequence
c	37	220.4	4.8	111803	2	AC136172	AC136172 Rattus no
c	38	188.8	4.1	213625	2	AC120950	AC120950 Rattus no
c	39	146.8	3.2	71599	2	AC139008	AC139008 Homo sapi
c	40	100.6	2.2	101	11	BV184269	BV184269 sqnm14402
c	41	94.2	2.0	374	10	AY091549S1	AY091549 Mus muscu
c	42	91	2.0	907	9	BC031942	BC031942 Homo sapi
c	43	89.8	2.0	707	6	CQ760483	CQ760483 Sequence
c	44	88.8	1.9	888	6	CQ760486	CQ760486 Sequence
c	45	88.8	1.9	900	6	CQ760482	CQ760482 Sequence

ALIGNMENTS

RESULT 1
LOCUS A66650 4596 bp DNA linear PAT 29-MAR-1999
DEFINITION Sequence 6 from Patent WO9737016.
ACCESSION A66650
VERSION A66650.1 GI:4538141
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4596)
Wallach,D., Malinin,N., Boldin,M., Kovalenko,A. and Mett,I.
MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR
PREPARATION AND USE
JOURNAL Patent: WO 9737016-A 6 09-OCT-1997;
YEDA RES & DEV (IL)
COMMENT Other publication AU 2175597 19971022.
FEATURES
Location/Qualifiers
source
1..4596
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 4596; DB 6; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGCGGGGGGACTGTGCCGTGTGGAACCTGTAGCTGTGTTGAAAGGTGGACTCTGTTACCATTTG 60
Db 1 AGCGGGGGGACTGTGCCGTGTGGAACCTGTAGCTGTGTTGAAAGGTGGACTCTGTTACCATTTG 60
Qy 61 AGGATGTTTGGAGGATGAGTATGTGTGGCAGGACACATTAACAGCGAGACCCCTTT 120
Db 61 AGGATGTTTGGAGGATGAGTATGTGTGGCAGGACACATTAACAGCGAGACCCCTTT 120
Qy 121 GCCCCTGCTTCTCCCCCAACCAAGCTGACCTGTGTTCTCCAGGTCTGGGATTTCTA 180
Db 121 GCCCCTGCTTCTCCCCCAACCAAGCTGACCTGTGTTCTCCAGGTCTGGGATTTCTA 180
Qy 181 AGTGACCTGCTGTGTGTTGGTCTCTCTCAGGATGAGCAACAGCTGGAGATGGCAGTG 240
Db 181 AGTGACCTGCTGTGTGTTGGTCTCTCTCAGGATGAGCAACAGCTGGAGATGGCAGTG 240

Db 181 AGTGACCTGCTCTGTGTTGGTCTCTCTCAGGATGAGCACAAAGCTGGAGATGGCAGTG 240
QY 241 ATGGAAATGGCTGCTGCCAGTGCCTCCCTGGCTCAGCAGTGGGCGACGACGAAGAACTCCCC 300
Db 241 ATGGAAATGGCTGCTGCCAGTGCCTCCCTGGCTCAGCAGTGGGCGACGACGAAGAACTCCCC 300
QY 301 AAGCCAAAGGAGAAAGACGCGCCACTCTGGGGAAGAAACAGAGCTCCGTCTTACAAAGCTTGAG 360
Db 301 AAGCCAAAGGAGAAAGACGCGCCACTCTGGGGAAGAAACAGAGCTCCGTCTTACAAAGCTTGAG 360
QY 361 GCCGTGGGAAGAGCCCTCTGTCTTCGCGAAAGTGGGAGATCCTGAATGAAGTGAATACC 420
Db 361 GCCGTGGGAAGAGCCCTCTGTCTTCGCGAAAGTGGGAGATCCTGAATGAAGTGAATACC 420
QY 421 AAGGGCAGACCAAGGAAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATCATCCGCCAG 480
Db 421 AAGGGCAGACCAAGGAAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATCATCCGCCAG 480
QY 481 GCTGAGTGTGAGATAGCCAAAGATTACGCCCACTTTTCAGAACGGCATTTTCATCGCT 540
Db 481 GCTGAGTGTGAGATAGCCAAAGATTACGCCCACTTTTCAGAACGGCATTTTCATCGCT 540
QY 541 GGGTCCAAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCAAACAATGTGCCCAT 600
Db 541 GGGTCCAAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCAAACAATGTGCCCAT 600
QY 601 GCTACAGAGGCAAAATGSCCGTGTGTGTTGGAGGGAAGCGTCGACGAAGCCCGG 660
Db 601 GCTACAGAGGCAAAATGSCCGTGTGTGTTGGAGGGAAGCGTCGACGAAGCCCGG 660
QY 661 AAGAAACGGAAGAAAGAGCTCAAAGTCCCTGGCTCATGCAAGGAGTGGCTTGGCCAAA 720
Db 661 AAGAAACGGAAGAAAGAGCTCAAAGTCCCTGGCTCATGCAAGGAGTGGCTTGGCCAAA 720
QY 721 CCCCTCCCAGGACCCCTGAGCAGGAGAGTGACCAATCCAGTGTGAGGAGATGAGTCT 780
Db 721 CCCCTCCCAGGACCCCTGAGCAGGAGAGTGACCAATCCAGTGTGAGGAGATGAGTCT 780
QY 781 CCACTCGGCGCCCATATGTTAGAAACACCCGCGAGTTCAACAAGCTCTGAAGGAACCA 840
Db 781 CCACTCGGCGCCCATATGTTAGAAACACCCGCGAGTTCAACAAGCTCTGAAGGAACCA 840
QY 841 GGCCTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGCTACGGCGGCTCTGCTCGA 900
Db 841 GGCCTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGCTACGGCGGCTCTGCTCGA 900
QY 901 TCAGAACTCCAAAACAGTACAGCCCTTGGCAATGTCTGAACCAAGTGTGGAACCTGCAC 960
Db 901 TCAGAACTCCAAAACAGTACAGCCCTTGGCAATGTCTGAACCAAGTGTGGAACCTGCAC 960
QY 961 CACCCCGAGAGCGAGGCCCCCTGCCCCCTGCCACGACCCCTTCCCTATAGCAGACTG 1020
Db 961 CACCCCGAGAGCGAGGCCCCCTGCCCCCTGCCACGACCCCTTCCCTATAGCAGACTG 1020
QY 1021 CCTCATCCCTTCCCATTTCAACCTCTCCAGCCCTGGAAACCTCACCTCTGAGTCTCTTC 1080
Db 1021 CCTCATCCCTTCCCATTTCAACCTCTCCAGCCCTGGAAACCTCACCTCTGAGTCTCTTC 1080
QY 1081 CTGGGCAAACTGGCTGTGTGAGACAGCAGGAACCTTGGCTGACCCACACCTGAGCAAA 1140
Db 1081 CTGGGCAAACTGGCTGTGTGAGACAGCAGGAACCTTGGCTGACCCACACCTGAGCAAA 1140
QY 1141 CTGGCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCCTG 1200
Db 1141 CTGGCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCCTG 1200
QY 1201 TCTCGTGGTCCCATGAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGC 1260
Db 1201 TCTCGTGGTCCCATGAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGC 1260
QY 1261 AGCGTGAAGCTCAAGCCAGGCCCCAGCCTGACAGCCTGCGCAAGCCTGGGCGACCG 1320
Db 1261 AGCGTGAAGCTCAAGCCAGGCCCCAGCCTGACAGCCTGCGCAAGCCTGGGCGACCG 1320

QY 1321 GGCTCCAGATCCCGGAGCCAGCCCCAAAACCTGAGACAAACGAGGGTGTCTTGCTCACT 1380
Db 1321 GGCTCCAGATCCCGGAGCCAGCCCCAAAACCTGAGACAAACGAGGGTGTCTTGCTCACT 1380
QY 1381 GAGAAACTCAAGCCAGTGGATTATAGTACCGAGAAAGATGTCTCACTGGGCGACGCCACAG 1440
Db 1381 GAGAAACTCAAGCCAGTGGATTATAGTACCGAGAAAGATGTCTCACTGGGCGACGCCACAG 1440
QY 1441 CTCGCTGGGCGAGAGCTCTTCGAGAGGTGCACAGGATGAGAGCAAGACAGCTGGC 1500
Db 1441 CTCGCTGGGCGAGAGCTCTTCGAGAGGTGCACAGGATGAGAGCAAGACAGCTGGC 1500
QY 1501 TTCCAGTGCCTGTCAAAAAGGTCGCTCGCTGGAAGTATTTTCGGGCGAGAGGCTGATGCA 1560
Db 1501 TTCCAGTGCCTGTCAAAAAGGTCGCTCGCTGGAAGTATTTTCGGGCGAGAGGCTGATGCA 1560
QY 1561 TGTGAGGATTGACCTCAACCCAGAAATTTGCTCCCTTTGTATGGAGCTGTGAGAAAGGCT 1620
Db 1561 TGTGAGGATTGACCTCAACCCAGAAATTTGCTCCCTTTGTATGGAGCTGTGAGAAAGGCT 1620
QY 1621 TGGGTCAACATCTTCATGGAGCTGCTGGAAGTGGCTCCCTGGGCGAGCTGGTCAAGGAG 1680
Db 1621 TGGGTCAACATCTTCATGGAGCTGCTGGAAGTGGCTCCCTGGGCGAGCTGGTCAAGGAG 1680
QY 1681 CAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCGAGGCTGGAGGCTGTG 1740
Db 1681 CAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCGAGGCTGGAGGCTGTG 1740
QY 1741 GAATACCTCCACTCAAGAAAGTCTGCAATGGGACCTCAAAAGCTGACAAAGCTGCTCTG 1800
Db 1741 GAATACCTCCACTCAAGAAAGTCTGCAATGGGACCTCAAAAGCTGACAAAGCTGCTCTG 1800
QY 1801 TCCAGCATGAGGAGCCAGCAGCCCTGTGACATTTTGGCCATGTGTGTCTTCAACCT 1860
Db 1801 TCCAGCATGAGGAGCCAGCAGCCCTGTGACATTTTGGCCATGTGTGTCTTCAACCT 1860
QY 1861 GATGCTCTGGGAAAGTCTTGTCTCAAGGGAATACATCTCTGGCAAGAGCCCAATG 1920
Db 1861 GATGCTCTGGGAAAGTCTTGTCTCAAGGGAATACATCTCTGGCAAGAGCCCAATG 1920
QY 1921 GCTCGGAGTGTGTCTGGGCGAGAGCTGCAAGCAGAGTGGATGTCTGAGGAGCTGC 1980
Db 1921 GCTCGGAGTGTGTCTGGGCGAGAGCTGCAAGCAGAGTGGATGTCTGAGGAGCTGC 1980
QY 1981 TGTATGATGTGCAATGCTCAAGCGCTGCCACCCCTGGAATCTTCTCCGAGGCGC 2040
Db 1981 TGTATGATGTGCAATGCTCAAGCGCTGCCACCCCTGGAATCTTCTCCGAGGCGC 2040
QY 2041 TCTGCTCTCAAGATTGCCAGGAGCTCCGCTGTGAGGAGATCCCACTCTCTGCGCC 2100
Db 2041 TCTGCTCTCAAGATTGCCAGGAGCTCCGCTGTGAGGAGATCCCACTCTCTGCGCC 2100
QY 2101 CCTCTCAGCCCGAGGCTCCAGAGGCTGAGGAAAGAGCCCATCCACCGCTGTCT 2160
Db 2101 CCTCTCAGCCCGAGGCTCCAGAGGCTGAGGAAAGAGCCCATCCACCGCTGTCT 2160
QY 2161 GCAGCGAGCTGGGAGGAGGTGAAACCGGCACTACAGCAAGTGGAGGTCTGAAGAGC 2220
Db 2161 GCAGCGAGCTGGGAGGAGGTGAAACCGGCACTACAGCAAGTGGAGGTCTGAAGAGC 2220
QY 2221 CTTTGGAGGGGAGATATAAAGAACCAAGACATCAACCGCCAAATCAAGCCAAATACCAC 2280
Db 2221 CTTTGGAGGGGAGATATAAAGAACCAAGACATCAACCGCCAAATCAAGCCAAATACCAC 2280
QY 2281 CAGACCTCTCAGCCCGAGGAGCTTTCGCAAGGCGCCAGGCGCCCGGCGAGCT 2340
Db 2281 CAGACCTCTCAGCCCGAGGAGGAGCTTTCGCAAGGCGCCAGGCGCCCGGCGAGCT 2340
QY 2341 GAGGAGACAAAGGCGAGAGCCCTTAAGCTTCAAGCTCTCTCCCAACAGAGCCCCAGAG 2400
Db 2341 GAGGAGACAAAGGCGAGAGCCCTTAAGCTTCAAGCTCTCTCCCAACAGAGCCCCAGAG 2400

Df	4561	CAATAAAAGTTTGGGTGATGACAAATGTTAAAAA	4596
RESULT 2			
AX429240	LOCUS	AX429240	4596 bp DNA linear PAT 21-JUN-2002
AX429240	DEFINITION	Sequence 17 from Patent EP1201765.	
AX429240	ACCESSION	AX429240	
AX429240.1	VERSION	GI:21540550	
KEYWORDS		synthetic construct	
SOURCE		synthetic construct	
ORGANISM		other sequences; artificial sequences.	
REFERENCE		1.	
AUTHORS		Schubart,D., Habenberger,P., Stein-Gerlach,M. and Bevec,D.	
TITLE		Cellular kinases involved in cytomegalovirus infection and their inhibition	
JOURNAL		Patent: EP 1201765-A 17 02-MAY-2002;	
FEATURES		Axxima Pharmaceuticals Aktiensgesellschaft (DE)	
source		Location/Qualifiers	
		1..4596	
		/organism="synthetic construct"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:32630"	
		/note="N/A"	
ORIGIN			
	Query Match	99.6%; Score 4578.4; DB 6; Length 4596;	
	Best Local Similarity	100.0%; Pred. No. 0;	
	Matches 4590; Conservative	0; Mismatches 1; Indels 1; Gaps 1;	
Qy	5	GCGGACTGTCGCTGGGAACGTCTAGCTGTGA-AGGTGACTCTGTTACCATTGAGG	63
Df	5	GCGGACTGTCGCTGTGGAACTGTAGCTGTGAGAGTGAGCTCTGTTACCATTTGAGG	64
Qy	64	ATGTTTTGGAGTAGATGATGTGTGGCAGAGGCACATAAAAGGACGAGACCCCTTTGCC	123
Df	65	ATGTTTTGGAGTAGATGATGTGTGGCAGAGGCACATAAAAGGACGAGACCCCTTTGCC	124
Qy	124	CTTGCTTTCTCCCCAACCCAAGGCTGACCTGTGTTCTCCAGGTCTGGGATTTCTAAGT	183
Df	125	CCTGCTTTCTCCCCAACCCAAGGCTGACCTGTGTTCTCCAGGTCTGGGATTTCTAAGT	184
Qy	184	GACTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAAGCTGGGAGATGGCAGTGATG	243
Df	185	GACTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAAGCTGGGAGATGGCAGTGATG	244
Qy	244	GAATGCTCTGCCAGTGCCCTTGCTCAGCAGTGGGACGAGAGAACTCCCCAAAG	303
Df	245	GAATGCTCTGCCAGTGCCCTTGCTCAGCAGTGGGACGAGAGAACTCCCCAAAG	304
Qy	304	CCAAAGAGAGACGCGCCACCTGGGGAAGAAACAGAGCTCCGCTCTACAAGCTTGAGGCC	363
Df	305	CCAAAGAGAGACGCGCCACCTGGGGAAGAAACAGAGCTCCGCTCTACAAGCTTGAGGCC	364
Qy	364	GTGAGAAGAGCCCTGTGTTTCTCGGAAGATGGGAGATCCTGAATGACGTGATTAACCAAG	423
Df	365	GTGAGAAGAGCCCTGTGTTTCTCGGAAGATGGGAGATCCTGAATGACGTGATTAACCAAG	424
Qy	424	GGCACACCAAGNAGGCTCCGAGGACGGCCAGCTGCCATCTCTATCATCGCCAGGCT	483
Df	425	GGCACACCAAGNAGGCTCCGAGGACGGCCAGCTGCCATCTCTATCATCGCCAGGCT	484
Qy	484	GAGTGTGAGAATAGCCAAAGAGTTTCAGCCCCACCTTTTCAGAACGCAATTTTCATCGCTGGG	543
Df	485	GAGTGTGAGAATAGCCAAAGAGTTTCAGCCCCACCTTTTCAGAACGCAATTTTCATCGCTGGG	544
Qy	544	TCCAAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCACCAATGTGGCCCATGCT	603
Df	545	TCCAAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCACCAATGTGGCCCATGCT	604
Qy	604	ACAGAGGCAAAATGGGCCCTGTGTGTTGGAAAGGGAAGCGTCGACAGCAAAGCCCGGAAG	663

Qy	1744	TACCTCCACTCAGGAAGGATTCTGCATGGGACGCTCAAAGCTGCAACAGTGTCTCTGTCC	1803
Db	1745	TACCTCCACTCAGGAAGGATTCTGCATGGGACGCTCAAAGCTGCAACAGTGTCTCTGTCC	1804
Qy	1804	AGCATGGGAGCCACGACGCCCTCTGTGACATTTGGCCATGCTGTGTCTTCAACCTGAT	1863
Db	1805	AGCATGGGAGCCACGACGCCCTCTGTGACATTTGGCCATGCTGTGTCTTCAACCTGAT	1864
Qy	1864	GGCTTGGGAAAGTCTTGTCTCACAGGGACTACATCCCTGGCCACAGAGACCACATGGCT	1923
Db	1865	GGCTTGGGAAAGTCTTGTCTCACAGGGACTACATCCCTGGCCACAGAGACCACATGGCT	1924
Qy	1924	CCGAGGTGTGTCTGGGACAGAGCTGGGACGCCAAGGTGGATGTCTGGAGCAGCTGTGT	1983
Db	1925	CCGAGGTGTGTCTGGGACAGAGCTGGGACGCCAAGGTGGATGTCTGGAGCAGCTGTGT	1984
Qy	1984	ATGATGTGCACTGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGCGCGCTC	2043
Db	1985	ATGATGTGCACTGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGCGCGCTC	2044
Qy	2044	TGCCTCAAGATTGCCAGGAGCCTCGCCTCTGTGAGGAGATCCACACCTCTCTGCGCCCT	2103
Db	2045	TGCCTCAAGATTGCCAGGAGCCTCGCCTCTGTGAGGAGATCCACACCTCTCTGCGCCCT	2104
Qy	2104	CTCACAGCCCAAGGATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCGTGTCTGCA	2163
Db	2105	CTCACAGCCCAAGGATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCGTGTCTGCA	2164
Qy	2164	GCGAGCTGGAGGAGGTGAACGGGGCACTACAGCAAGTGGAGGTCTGAAGAGCCT	2223
Db	2165	GCGAGCTGGAGGAGGTGAACGGGGCACTACAGCAAGTGGAGGTCTGAAGAGCCT	2224
Qy	2224	TGGAGGGAGAAATATAAGAACCAAGACATCCACGCCCAATCAAGCCAAATTACACACAG	2283
Db	2225	TGGAGGGAGAAATATAAGAACCAAGACATCCACGCCCAATCAAGCCAAATTACACACAG	2284
Qy	2284	ACCTCTCATGCCACGCGAGAGAGCTTTTCGCAAGAGGGCCCAAGGGCCCGGCCACAGCTGAG	2343
Db	2285	ACCTCTCATGCCACGCGAGAGAGCTTTTCGCAAGAGGGCCCAAGGGCCCGGCCACAGCTGAG	2344
Qy	2344	GAGACACAGGACGAGCCCTTAAGTCTCAGCTCTCTCTTCCACAGAGCCCCACAGAGCCA	2403
Db	2345	GAGACACAGGACGAGCCCTTAAGTCTCAGCTCTCTCTTCCACAGAGCCCCACAGAGCCA	2404
Qy	2404	AACAAGTCTCTCTCTGACTTTTCGCAAGAGAGAGTCTGGGATGTGGACACCTTACCT	2463
Db	2405	AACAAGTCTCTCTCTGACTTTTCGCAAGAGAGAGTCTGGGATGTGGGAAACCTTACCT	2464
Qy	2464	CTGTCTCTCTGGAGCCAGCCCTTCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC	2523
Db	2465	CTGTCTCTCTGGAGCCAGCCCTTCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC	2524
Qy	2524	GTCCCGGACGAGAACTGCAGCAGCTGGAATATGAAATATTCCTCAACAGCCTGTCCAG	2583
Db	2525	GTCCCGGACGAGAACTGCAGCAGCTGGAATATGAAATATTCCTCAACAGCCTGTCCAG	2584
Qy	2584	CCATTTTCTCTGGAGGACGAGCAAAATTTCTCTGTGCTCTCAGCATGACAGCCTCTCC	2643
Db	2585	CCATTTTCTCTGGAGGACGAGCAAAATTTCTCTGTGCTCTCAGCATGACAGCCTCTCC	2644
Qy	2644	CTGTCTGATGACAGTGAAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGACACCCCTG	2703
Db	2645	CTGTCTGATGACAGTGAAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGACACCCCTG	2704
Qy	2704	AGCTCAGGCGTACACTCTCTGGAGCAGCCGAGGCTCGAAGCTCCAGCTGGAACATG	2763
Db	2705	AGCTCAGGCGTACACTCTCTGGAGCAGCCGAGGCTCGAAGCTCCAGCTGGAACATG	2764
Qy	2764	GTGCTGGCCCGGGGCGGCCACCGACACCCCAAGCTTATTTCAATGGTGTGAAGTCCAA	2823
Db	2765	GTGCTGGCCCGGGGCGGCCACCGACACCCCAAGCTTATTTCAATGGTGTGAAGTCCAA	2824

Qy	2824	ATACAGTCTCTTAATGTGGAACACCTGCATCCGGAGTTTCCACCGGTTCAAAAGTGGGA	2883
Db	2825	ATACAGTCTCTTAATGTGGAACACCTGCATCCGGAGTTTCCACCGGTTCAAAAGTGGGA	2884
Qy	2884	GACATCCGCCACTGGCATCAGCAGCCAGATCCCAAGCTTCAGCTTGGTGGTCAACAA	2943
Db	2885	GACATCCGCCACTGGCATCAGCAGCCAGATCCCAAGCTTCAGCTTGGTGGTCAACAA	2944
Qy	2944	GACGGGAGCCTGTTCCTCTACGACATGGAGGTGCAGATCGGGCATCGACTGCAGTGC	3003
Db	2945	GACGGGAGCCTGTTCCTCTACGACATGGAGGTGCAGATCGGGCATCGACTGCAGTGC	3004
Qy	3004	ACACTGCGCCCTGATGACGACTTCCCTCTGAGCTGGAGGTTCAAGCATGGCAGCTGAG	3063
Db	3005	ACACTGCGCCCTGATGACGACTTCCCTCTGAGCTGGAGGTTCAAGCATGGCAGCTGAG	3064
Qy	3064	AACAGGCCCTAAACCTCGCCCTCCACCGCGCTCCACACTGCCGGAAGCAGCTTCCCTG	3123
Db	3065	AACAGGCCCTAAACCTCGCCCTCCACCGCGCTCCACACTGCCGGAAGCAGCTTCCCTG	3124
Qy	3124	CTCGGTGCAGATGCTGCCCTGAAAAACACAGGCTCAGCCGTTCCAGGGGATTGCCAGCC	3183
Db	3125	CTCGGTGCAGATGCTGCCCTGAAAAACACAGGCTCAGCCGTTCCAGGGGATTGCCAGCC	3184
Qy	3184	CCCCGGTCAAGTGGGAAACAGGGCTTCGACAGCAAGTGGGGGCAAGCAATGCC	3243
Db	3185	CCCCGGTCAAGTGGGAAACAGGGCTTCGACAGCAAGTGGGGGCAAGCAATGCC	3244
Qy	3244	TCCAGGATTTTCACTGAGCCCTGCCACCCCTGCTGAAAAACATCCCGCCAGTGAA	3303
Db	3245	TCCAGGATTTTCACTGAGCCCTGCCACCCCTGCTGAAAAACATCCCGCCAGTGAA	3304
Qy	3304	GAGACAGAGGAGTGGCAGGAGTTACTTGGGGAACAAAAACAGGGATCTTTTCTGCC	3363
Db	3305	GAGACAGAGGAGTGGCAGGAGTTACTTGGGGAACAAAAACAGGGATCTTTTCTGCC	3364
Qy	3364	CTGTCTCCAGTTCGAGTTGGCTGACCCGCTTGGATCAGTGCACATTTTGTGGCAGACAG	3423
Db	3365	CTGTCTCCAGTTCGAGTTGGCTGACCCGCTTGGATCAGTGCACATTTTGTGGCAGACAG	3424
Qy	3424	GGAGAGCAGCTTCCAGCTTGGTTCAGAAAGGGTGGGCGAGCCCTTCGCCCTCACCCCTC	3483
Db	3425	GGAGAGCAGCTTCCAGCTTGGTTCAGAAAGGGTGGGCGAGCCCTTCGCCCTCACCCCTC	3484
Qy	3484	CAGCTCTGTGAGAGTGTCAAGTGTCTAAGGGCCCAAACTCAGGTTTCAGTGCAGAACCA	3543
Db	3485	CAGCTCTGTGAGAGTGTCAAGTGTCTAAGGGCCCAAACTCAGGTTTCAGTGCAGAACCA	3544
Qy	3544	GGTCAGAGGTATGCCCCCGCTAGAGTTAAGGGGGCCCTCTAAACCCCTTGGCTGGCCTC	3603
Db	3545	GGTCAGAGGTATGCCCCCGCTAGAGTTAAGGGGGCCCTCTAAACCCCTTGGCTGGCCTC	3604
Qy	3604	ACCTGGCAGCTACCCCTTTTGGGTTCAGGGGAAAGAAATGCTGACCTTGGGAAAGGCT	3663
Db	3605	ACCTGGCAGCTACCCCTTTTGGGTTCAGGGGAAAGAAATGCTGACCTTGGGAAAGGCT	3664
Qy	3664	CCCTGGTAGAATAACACACTTTTTCAGGTTGTTCAGACACAGCTCTCAGTTCACCTCT	3723
Db	3665	CCCTGGTAGAATAACACACTTTTTCAGGTTGTTCAGACACAGCTCTCAGTTCACCTCT	3724
Qy	3724	GGTTTCAGCAAGGACCAAGAGGTGTGAAGTGAAGTGGTTTCTCAGTCCCCCAGACATGT	3783
Db	3725	GGTTTCAGCAAGGACCAAGAGGTGTGAAGTGAAGTGGTTTCTCAGTCCCCCAGACATGT	3784
Qy	3784	GCCCCCTTGTGTGGTACCACTTCTTCCAGAGCAGCAGGCCCCCAGGCCCCCTTCAGGC	3843
Db	3785	GCCCCCTTGTGTGGTACCACTTCTTCCAGAGCAGCAGGCCCCCAGGCCCCCTTCAGGC	3844
Qy	3844	CCAGCATGCCCCAGACTCGCTGCACTCAGTTCCTCATCTGTAAGGTTGAAGGTTGAT	3903
Db	3845	CCAGCATGCCCCAGACTCGCTGCACTCAGTTCCTCATCTGTAAGGTTGAAGGTTGAT	3904
Qy	3904	GCAGGATATGCTGCACAGGAAACAGTCTGTGGAATGGACATGATCAGTCTAAGGAAAGCAG	3963

QY 1084 GGCAAACTGGCCCTGTGTAGACGACGAAACCCCTTGGCTGTGACCCACACCTGAGCAAACTG 1143
DB 1085 GGCAAACTGGCCCTGTGTAGACGACGAAACCCCTTGGCTGTGACCCACACCTGAGCAAACTG 1144
QY 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGTCT 1203
DB 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGTCT 1204
QY 1204 CGTGGTCCCATGAGAACTTTTCTGTGGAGGAATACCTTAGTGCATGTCTTCAAGGCAGC 1263
DB 1205 CGTGGTCCCATGAGAACTTTTCTGTGGAGGAATACCTTAGTGCATGTCTTCAAGGCAGC 1264
QY 1264 GTGAGCTCAAGCCAGGCCCCACAGCTGTACAGCTGTGGCCAAAGACTTGGGAGCAAGGGGC 1323
DB 1265 GTGAGCTCAAGCCAGGCCCCACAGCTGTACAGCTGTGGCCAAAGACTTGGGAGCAAGGGGC 1324
QY 1324 TCCAGATCCCGGGAGCCCCAGCCCAAAACTGAGGACACAGAGGTGCTCTCACTGAG 1383
DB 1325 TCCAGATCCCGGGAGCCCCAGCCCAAAACTGAGGACACAGAGGTGCTCTCACTGAG 1384
QY 1384 AAACCTCAAGCCAGTGGATTATAGATTACCGAGAAAGTCCACTGGGCCACGACACAGCTC 1443
DB 1385 AAACCTCAAGCCAGTGGATTATAGATTACCGAGAAAGTCCACTGGGCCACGACACAGCTC 1444
QY 1444 CGCCTGGGACAGAGCTCCTTCGAGAGGTGCAAGGATGGAGGAACAAGCACTGGCTTC 1503
DB 1445 CGCCTGGGACAGAGCTCCTTCGAGAGGTGCAAGGATGGAGGAACAAGCACTGGCTTC 1504
QY 1504 CAGTGGCTGTCAAAAGGTGGCTGGAAGTATTTGGGGCAGAGGAGCTGATGGCATGT 1563
DB 1505 CAGTGGCTGTCAAAAGGTGGCTGGAAGTATTTGGGGCAGAGGAGCTGATGGCATGT 1564
QY 1564 GCAGGATTGACCTCACCCAGAAATTTCCCTTTGTATGGAGCTGTGAGAAAGGCCCTTGG 1623
DB 1565 GCAGGATTGACCTCACCCAGAAATTTCCCTTTGTATGGAGCTGTGAGAAAGGCCCTTGG 1624
QY 1624 GTCAACATCTTTCATGGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGGAGAG 1683
DB 1625 GTCAACATCTTTCATGGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGGAGAG 1684
QY 1684 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGCCCTGGAGGCTTGGAA 1743
DB 1685 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGCCCTGGAGGCTTGGAA 1744
QY 1744 TACCTCACTCACGAAGGATTTGCAATGGGACGTCGAAAGCTGCAACGCTGCTCTGTGCC 1803
DB 1745 TACCTCACTCACGAAGGATTTGCAATGGGACGTCGAAAGCTGCAACGCTGCTCTGTGCC 1804
QY 1804 AGCGATGGAGCCACGACGCCCTCTGTGACTTTGGCCATGTGTGTCTTTCAACCTGAT 1863
DB 1805 AGCGATGGAGCCACGACGCCCTCTGTGACTTTGGCCATGTGTGTCTTTCAACCTGAT 1864
QY 1864 GGCTTGGAAAGTCTTGTCTACAGGGGACTACATCCCTGGCCACAGAGCCCATGGCT 1923
DB 1865 GGCTTGGAAAGTCTTGTCTACAGGGGACTACATCCCTGGCCACAGAGCCCATGGCT 1924
QY 1924 CCGAGGTGGTGTGGGACAGGAGTGCAGACCCAAAGTGGATGTCTGGAGCAGCTGTGT 1983
DB 1925 CCGAGGTGGTGTGGGACAGGAGTGCAGACCCAAAGTGGATGTCTGGAGCAGCTGTGT 1984
QY 1984 ATGATGTGCAATGTCTCAAGCGGTGCCACCCCTGGGACTCAGTTCTTCCGAGGGCCGCTC 2043
DB 1985 ATGATGTGCAATGTCTCAAGCGGTGCCACCCCTGGGACTCAGTTCTTCCGAGGGCCGCTC 2044
QY 2044 TGCCTCAAGATTGCGAGGAGCTCCGCTGTGAGGAGATCCACCCCTCTGGGCCCT 2103
DB 2045 TGCCTCAAGATTGCGAGGAGCTCCGCTGTGAGGAGATCCACCCCTCTGGGCCCT 2104
QY 2104 CTCAAGCCAGGCGCATCCAGAGGGGTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2163
DB 2105 CTCAAGCCAGGCGCATCCAGAGGGGTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2164

QY 2164 GCGAGCTGGAGGGAAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAAGGCCCT 2223
DB 2165 GCGAGCTGGAGGGAAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAAGGCCCT 2224
QY 2224 TGGAGGGGAGAAATAAAGAACCAAGACATCCACCGCAAAATCAAGCAAATTTACCACCAG 2283
DB 2225 TGGAGGGGAGAAATAAAGAACCAAGACATCCACCGCAAAATCAAGCAAATTTACCACCAG 2284
QY 2284 ACCCTCCATGCCAGCCGAGAGAGCTTTGCGCAAGGGCCCCCAGGGCCCCCGGCGAGCTGAG 2343
DB 2285 ACCCTCCATGCCAGCCGAGAGAGCTTTGCGCAAGGGCCCCCAGGGCCCCCGGCGAGCTGAG 2344
QY 2344 GAGACAAGGAGAGAGCCCTTAAGCTCAGGCTCTCTCCACAGAGCCGCCAGAGGCA 2403
DB 2345 GAGACAAGGAGAGAGCCCTTAAGCTCAGGCTCTCTCTCCACAGAGCCGCCAGAGGCA 2404
QY 2404 AACAAAGTCTCTCTCTTGACTTTTTCAGCAAGGAGAGTCTGGGATGTGGGAAACCCCTTACT 2463
DB 2405 AACAAAGTCTCTCTCTTGACTTTTTCAGCAAGGAGAGTCTGGGATGTGGGAAACCCCTTACT 2464
QY 2464 CTGTCTCTCTCTGAGCCAGCCCTCTGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2523
DB 2465 CTGTCTCTCTCTGAGCCAGCCCTCTGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2524
QY 2524 GTCCCGAGAGAGAACTGCGAGAGCTGGAATAAGATTTATTCCTCAACAGGCTGTCCAG 2583
DB 2525 GTCCCGAGAGAGAACTGCGAGAGCTGGAATAAGATTTATTCCTCAACAGGCTGTCCAG 2584
QY 2584 CCATTTTCTCTGGAGGAGAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCTCTCC 2643
DB 2585 CCATTTTCTCTGGAGGAGAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCTCTCC 2644
QY 2644 CTGTCCGATGACAGTGAAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGAGCACCCCTG 2703
DB 2645 CTGTCCGATGACAGTGAAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGAGCACCCCTG 2704
QY 2704 AGCTCAGGCTACTCTCTGGAGCAGCCAGGCTCGAAGCTCGAGCTGGAACAATG 2763
DB 2705 AGCTCAGGCTACTCTCTGGAGCAGCCAGGCTCGAAGCTCGAGCTGGAACAATG 2764
QY 2764 GTGTGCCCCGGGGGGCCACCGACACCCCAAGCTATTTCATGTTGTAAGTCCAA 2823
DB 2765 GTGTGCCCCGGGGGGCCACCGACACCCCAAGCTATTTCATGTTGTAAGTCCAA 2824
QY 2824 ATACAGTCTCTTAATGTGAACACCTGCACATCCGGAGTTCCACCGGTTCAAGTGGGA 2883
DB 2825 ATACAGTCTCTTAATGTGAACACCTGCACATCCGGAGTTCCACCGGTTCAAGTGGGA 2884
QY 2884 GACATGCCACTGGCATCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCAACCAA 2943
DB 2885 GACATGCCACTGGCATCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCAACCAA 2944
QY 2944 GACGGGAGCTGTTGCTTACGATGAGGTGCCAGACTCGGGCATCGACTCGAGTGC 3003
DB 2945 GACGGGAGCTGTTGCTTACGATGAGGTGCCAGACTCGGGCATCGACTCGAGTGC 3004
QY 3004 ACCTGCCCCCTGATGTCAGCTTCGCTGAGCTGGAGGGTCAAGCATGGCAGCTGAG 3063
DB 3005 ACCTGCCCCCTGATGTCAGCTTCGCTGAGCTGGAGGGTCAAGCATGGCAGCTGAG 3064
QY 3064 AACAGGCCCTTAAACCTTCCCTTCAACCGCCGCTCCACACTGCGGAAAGAGCCTTCTG 3123
DB 3065 AACAGGCCCTTAAACCTTCCCTTCAACCGCCGCTCCACACTGCGGAAAGAGCCTTCTG 3124
QY 3124 CTGGTGCAGATGCTGCTTGAATAACAGAGGCTCAGCTTCCAGGGATTTGCCAGCC 3183
DB 3125 CTGGTGCAGATGCTGCTTGAATAACAGAGGCTCAGCTTCCAGGGATTTGCCAGCC 3184
QY 3184 CCCCGGCTCAGTGGGAAACAGGGCTTCGAGCAGCAAGTGGGGCAAGAGCAATGCC 3243
DB 3185 CCCCGGCTCAGTGGGAAACAGGGCTTCGAGCAGCAAGTGGGGCAAGAGCAATGCC 3244
QY 3244 TCCAGGATTTACACCTGAGCCCTGCCCAACCCCTGTCTGTAATAAATCCGCCACGTGAA 3303

Db 3245 TCCAGGATTTTCAACCTGAGCCCTGCCCCACCTCTGCTGAATAACATCCGCCAGTGAA 3304
QY 3304 GAGACAGAGGAGGATGCGAGAGTTACTCTGGGAAACAAAACAGGAGTCTTTTCTGCC 3363
Db 3305 GAGACAGAGGAGGATGCGAGAGTTACTCTGGGAAACAAAACAGGAGTCTTTTCTGCC 3364
QY 3364 CTTGCTCAGATCGATTGGCTTGACCCGTTGGATCAGTGACCAATTTTGTGGCAGACAGG 3423
Db 3365 CTTGCTCAGATCGATTGGCTTGACCCGTTGGATCAGTGACCAATTTTGTGGCAGACAGG 3424
QY 3424 GGAGAGACGTTTCCAGCTGGGTGAGAGGGGTGGCGAGCCCTTCGCGCCCTCACCCCTC 3483
Db 3425 GGAGAGACGTTTCCAGCTGGGTGAGAGGGGTGGCGAGCCCTTCGCGCCCTCACCCCTC 3484
QY 3484 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGCCCCAAACTCAGTTCAGTGCAAGACCA 3543
Db 3485 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGCCCCAAACTCAGTTCAGTGCAAGACCA 3544
QY 3544 GGTACAGAGTATGCCCGCCGTAGGTTAAGGGGGCCCTCTAAACCCCTTGCTTGGCTGCCTC 3603
Db 3545 GGTACAGAGTATGCCCGCCGTAGGTTAAGGGGGCCCTCTAAACCCCTTGCTTGGCTGCCTC 3604
QY 3604 ACTGSCCAGCTCACCCCTTTTGGGTGAGGGGAAAAGAAATGCCTGACCCCTGGGAAGGCT 3663
Db 3605 ACTGSCCAGCTCACCCCTTTTGGGTGAGGGGAAAAGAAATGCCTGACCCCTGGGAAGGCT 3664
QY 3664 CCTGCTAGAAATACACCACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCT 3723
Db 3665 CCTGCTAGAAATACACCACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCT 3724
QY 3724 GGTTCAGCAAGACCAAGAGGTGTGTAAGTGAAGTGTCTCAGTCCCCAGACATGT 3783
Db 3725 GGTTCAGCAAGACCAAGAGGTGTGTAAGTGAAGTGTCTCAGTCCCCAGACATGT 3784
QY 3784 GCGCTTCTGCTGCTACCACTCTTCCAGAGCAGCAGCGCCCGAGCCCTTCAGGC 3843
Db 3785 GCGCTTCTGCTGCTACCACTCTTCCAGAGCAGCAGCGCCCGAGCCCTTCAGGC 3844
QY 3844 CCAGCACTCCCCAGACTCGCTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGTGAT 3903
Db 3845 CCAGCACTCCCCAGACTCGCTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGTGAT 3904
QY 3904 GAGGATATGCTCAGAGAACAGTCTGGATGAGCATGATCAGTCTAAGGAAGCAG 3963
Db 3905 GCAGGATATGCTCAGAGAACAGTCTGGATGAGCATGATCAGTCTAAGGAAGCAG 3964
QY 3964 CAGAGAGACGTCCCGCGCCCGAGCCCACTATCAGTGTCCAGCGTGTGGTTCCCCAG 4023
Db 3965 CAGAGAGACGTCCCGCGCCCGAGCCCACTATCAGTGTCCAGCGTGTGGTTCCCCAG 4024
QY 4024 AGCAGAGCTCAGCATCACACTGACATCACTACCTGCTGCCCTGGCCAGAGGTAAGTCTGC 4083
Db 4025 AGCAGAGCTCAGCATCACACTGACATCACTACCTGCTGCCCTGGCCAGAGGTAAGTCTGC 4084
QY 4084 CGAGCGACTTTTGCACTCTGATGACCTCAAGACATTTTATGCTGCCCTCTGGCAGGGC 4143
Db 4085 CGAGCGACTTTTGCACTCTGATGACCTCAAGACATTTTATGCTGCCCTCTGGCAGGGC 4144
QY 4144 AGGGCAGGGCAGTGACATGTTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGACACA 4203
Db 4145 AGGGCAGGGCAGTGACATGTTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGACACA 4204
QY 4204 GTCTTGAGCTGTCCAGATGATGATCTCTTCAAACTCTTCCAGATTTCTCTAAGAAATA 4263
Db 4205 GTCTTGAGCTGTCCAGATGATGATCTCTTCAAACTCTTCCAGATTTCTCTAAGAAATA 4264
QY 4264 GCACCCCTTCCCATTTGCCAGCTTAGCTCTTCCAGGGGAGTACTCAGGACTC 4323
Db 4265 GCACCCCTTCCCATTTGCCAGCTTAGCTCTTCCAGGGGAGTACTCAGGACTC 4324
QY 4324 ACCTAGCATTAATCAGCTGTGAATCGTCAGGGGGTGTCTGCTAGCGCTCAACCTCTCTGGG 4383

Db 4325 ACGTAGCATTAATTCAGCTGTGAATCGTCAGGGGGTGTCTGCTAGCCTCAACCTCTGGG 4384
QY 4384 GCAGGGACGCCGAGAGCTCCGTGGGAGAAGCTATCCCAATCTTCCCAAGACAGAGCTT 4443
Db 4385 GCAGGGACGCCGAGAGCTCCGTGGGAGAAGCTATCCCAATCTTCCCAAGACAGAGCTT 4444
QY 4444 TGTCCAGCTGTCCACATTTGAGTCTGCTCCCGGGGAGAGAGCCCGGCCCCAGCAC 4503
Db 4445 TGTCCAGCTGTCCACATTTGAGTCTGCTCCCGGGGAGAGAGCCCGGCCCCAGCAC 4504
QY 4504 ATAAAGAACTGCAGCTTGGTCTGACAGAGTCTGGGTGTAGAGAACTCTTTGTAAAGCAA 4563
Db 4505 ATAAAGAACTGCAGCTTGGTCTGACAGAGTCTGGGTGTAGAGAACTCTTTGTAAAGCAA 4564
QY 4564 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4595
Db 4565 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4596

RESULT 4
HSDIK HSDIK 4596 bp mRNA linear PRI 13-JAN-1998
LOCUS H.sapiens mRNA for serine/threonine protein kinase, NIK.
DEFINITION Y10256
ACCESSION Y10256
VERSION Y10256.1 GI:1841433
KEYWORDS MAP kinase; NIK protein; serine/threonine protein kinase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Malinin,N.L., Boldin,M.P., Kovalenko,A.V. and Wallach,D.
TITLE MAP3K-related kinase involved in NF-kappaB induction by TNF, CD95 and IL-1
JOURNAL Nature 385 (6616), 540-544 (1997)
MEDLINE 97172277
PUBMED 9020361
REFERENCE 2 (bases 1 to 4596)
AUTHORS Wallach,D.
TITLE Direct Submission
JOURNAL Submitted (23-DEC-1996) D. Wallach, The Weizmann Institute, Dept of Membrane Research & Biophysics, Rehovot 76100, ISRAEL
COMMENT NIK is a serine/threonine protein-kinase, resembling several MAP kinase kinase kinases (MAP3K), that binds specifically to TRAF2, an adapter proteins associated, either directly or through interaction with other adapter proteins, with several receptors of the TNF/NGF family. NIK overexpression in cells activates the transcription factor NF kappa B. Cellular expression of kinase-deficient NIK-mutants blocks NF kappa B induction by TNF, by either of the two TNF receptors, by CD95 (Fas/Apo-1) and by TRADD, RIP and MORT1/FADD, adapter proteins that bind to these receptors. It also blocks NF kappa B induction by IL-1.

FEATURES

source
1..4596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="B cells"

misc_feature

137..149
/note="lysine-rich region"
233..3076

CDS

/codon_start=1
/product="NIK, serine/threonine protein-kinase"
/protein_id="CAA71306.1"
/db_xref="GI:1841434"
/db_xref="GOA:Q99558"
/db_xref="UniProt/Swiss-Prot:Q99558"

/translation="MAVMEACPGAGSAGVQKSLPKPKETPLGKKQSSVYKLEA
VEKSPVFCGEILNDVITTKGTEGSEAGAPAAISIIAQACENQSEFPTFSERIFI
AGSKQYSSQESLDQIPNNVAHATEGMARCVKWKRRKRSKSKSLAHAGVA
LAKPLPRPEQESCTIPVQDESPGAPVVRNTPQTKLPKPEGLQGLCKOLGEGLR
PALPKSLHLKLISPLQCLNHVWLHPDGGPLPLPTHPFPYSRLPHPPPLQPLQPK
PPLPSFLGKLKACVDSQKPLPDPHLSKLKACVDSKPLPLGPHLEPSCLSRGAHEKFSVE

EYLVAHLOGSVSSQAHSLTSLAKTWAARGSRREPSKPTEDNEGVLLTEKLKPVVYE
YREYVWATHQRLGRFGEVHREDKQTFQCAVKVRLEVPRAELMACAGLTSP
RIVLYGAVREGPVNTFMELLEGSLQVLKEQCLPEDRALYVLOALGLEYLSP
RRIHGDVADNVLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDIPTGETHMAPE
VVLGRSDAKVDVSSCMMLHMLNGCHPWTOFFRGPCLCKIASEPVPREIPPSCAP
LTAQRIQEGLEKEPIHVRUSAELGKVNRLQVVGKLSWPRGKYKEPRHPPPNOAY
EPTLHAQPRELSPRAGPRPABETVGRAPKLOPLPEPPPEPNKSPPLTUSKEESGMW
BPLFLSLPEAPNPSPERKATVPEQLOLEIEFLNLSQFPLEQEQEILSL
SIDLSLSDSEKNPSKASQSRDTLSGSHVSSWQAEARSSMMVLAIRGRPTDPS
YFNGVKVQIOISLNGEHLHIREFHRVKVGDIAIGISSQIPAAAFSLVTKDGPVRYDME
VPDSGIDLQCTLAPDGSFANWVKHGLENRP"

misc_feature

/note=".53

misc_feature

/note="kinase domain"

711..729

/note="proline-rich region"

ORIGIN

Query Match 99.6%; Score 4578.4; DB 9; Length 4596;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy	5	GGGGAGCTGTCGGGTGGAACGTGTAGCTGTGA-AGGTGGACTCTGTTACCAATTGAGG	63
Db			
Qy	5	GGGGAGCTGTCGGGTGGAACGTGTAGCTGTGAGAGTGGACTCTGTTACCAATTGAGG	64
Db			
Qy	64	ATGTTTGAGGATGAGTATGTGTGGCAGAGGCACACATAAAGCAGCAGACCCCTTTGCC	123
Db			
Qy	124	CCTGCTTTTCCCCAAACCCAGGCTGACCTGTGTTCTCCAGGCTCGGATTTCAAGT	183
Db			
Qy	125	CCTGCTTTTCCCCAAACCCAGGCTGACCTGTGTTCTCCAGGCTCGGATTTCAAGT	184
Db			
Qy	184	GACCTGCTCTGTGTTGTGCTCTCTCAGGATGACACAGCTCGGAGATGGCAGTGTG	243
Db			
Qy	185	GACCTGCTCTGTGTTGTGCTCTCTCAGGATGACACAGCTCGGAGATGGCAGTGTG	244
Db			
Qy	244	GAATATGGCTTCCCAAGTGCCCTGGCTCAGCAGTGGGCGACAGAGGAACCTCCCAAG	303
Db			
Qy	245	GAATATGGCTTCCCAAGTGCCCTGGCTCAGCAGTGGGCGACAGAGGAACCTCCCAAG	304
Db			
Qy	304	CCAAAGGAGAGAGACCGCCCACTGGGGAAGAAACAGAGCTCCGTCTCAAGCTTGAGCC	363
Db			
Qy	305	CCAAAGGAGAGAGACCGCCCACTGGGGAAGAAACAGAGCTCCGTCTCAAGCTTGAGCC	364
Db			
Qy	364	GTGAGAGAGCCCTGTGTTCTCGGAAAGTGGGAGATCCCTGAATGACGTGATACCAAG	423
Db			
Qy	365	GTGAGAGAGAGCCCTGTGTTCTCGGAAAGTGGGAGATCCCTGAATGACGTGATACCAAG	424
Db			
Qy	424	GGCAGACCAAGGAAGGCTCCGAGGCGAGGCGAGCTGCCATCTCTATCATCGCCAGGCT	483
Db			
Qy	425	GGCAGACCAAGGAAGGCTCCGAGGCGAGGCGAGCTGCCATCTCTATCATCGCCAGGCT	484
Db			
Qy	484	GAGTGTGAGAAATAGCAAGATTGAGCCCACTTTTCAGAACGCAATTTTCATCGCTGGG	543
Db			
Qy	485	GAGTGTGAGAAATAGCAAGATTGAGCCCACTTTTCAGAACGCAATTTTCATCGCTGGG	544
Db			
Qy	544	TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACATGTGGCCCATGCT	603
Db			
Qy	545	TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACATGTGGCCCATGCT	604
Db			
Qy	604	ACAGAGGCAAAATGGGCCGTGTGTGTTGGAAGGGAAGCGTCGACGAAAGCCCGGAAG	663
Db			
Qy	605	ACAGAGGCAAAATGGGCCGTGTGTGTTGGAAGGGAAGCGTCGACGAAAGCCCGGAAG	664
Db			
Qy	664	AAACGGAAGAAAGAGAGCTCAAAGTCCCTGGCTCATGCAAGAGTGGCTTTGGCCAAACCC	723
Db			
Qy	665	AAACGGAAGAAAGAGAGCTCAAAGTCCCTGGCTCATGCAAGAGTGGCTTTGGCCAAACCC	724
Db			
Qy	724	CTCCCGAGGACCCCTGAGCAGGAGAGTGCACCATCCAGTGCAGGAGATGATCTTCCA	783
Db			
Qy	725	CTCCCGAGGACCCCTGAGCAGGAGAGTGCACCATCCAGTGCAGGAGATGATCTTCCA	784
Db			

Qy	784	CTCGGCGCCCATATGTTAGAAACACACCCCGCAGTTTCAACAGCCTCTGTAAAGAAACAGGC	843
Db			
Qy	785	CTCGGCGCCCATATGTTAGAAACACACCCCGCAGTTTCAACAGCCTCTGTAAAGAAACAGGC	844
Db			
Qy	844	CTTGGGCAACTCTGTTTTTAAGCAGCTTGGCGAGGGCCTACGGCCGGCTCTCGCTCGATCA	903
Db			
Qy	845	CTTGGGCAACTCTGTTTTTAAGCAGCTTGGCGAGGGCCTACGGCCGGCTCTCGCTCGATCA	904
Db			
Qy	904	GAACTCCACAAACTGATTCAGCCCTTGCATATGCTGAAACACAGTGTGGAACCTGCACCAC	963
Db			
Qy	905	GAACTCCACAAACTGATTCAGCCCTTGCATATGCTGAAACACAGTGTGGAACCTGCACCAC	964
Db			
Qy	964	CCCAGAGCGAGGAGCCCTTGCCTGCCACGCAACCCCTTCCCTTATAGCAGACTGCTCTG	1023
Db			
Qy	965	CCCAGAGCGAGGAGCCCTTGCCTGCCACGCAACCCCTTCCCTTATAGCAGACTGCTCTG	1024
Db			
Qy	1024	CATCCCTTCCCATTCACACCTCTCCAGCCCTGGAAACCTCAACCTCTGGAGCTCTTCTG	1083
Db			
Qy	1025	CATCCCTTCCCATTCACACCTCTCCAGCCCTGGAAACCTCAACCTCTGGAGCTCTTCTG	1084
Db			
Qy	1084	GGCAAACTGGCCTGTGTAGACAGCCAGAAACCCCTTGCCTGACCCACACCTCAGCAAACTG	1143
Db			
Qy	1085	GGCAAACTGGCCTGTGTAGACAGCCAGAAACCCCTTGCCTGACCCACACCTCAGCAAACTG	1144
Db			
Qy	1144	GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTGGAGCCAGCTGCTGTCT	1203
Db			
Qy	1145	GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTGGAGCCAGCTGCTGTCT	1204
Db			
Qy	1204	CGTGTGCCATGAGAAATTTTCTGTGGAGGAATACCTAGTGTGATGTCTGTCAAGGAGC	1263
Db			
Qy	1205	CGTGTGCCATGAGAAATTTTCTGTGGAGGAATACCTAGTGTGATGTCTGTCAAGGAGC	1264
Db			
Qy	1264	GTGAGCTCAAGCCAGGCGCCACAGCTGCACAGCCTGSCAAAGACCTGGGAGCAGCGGGC	1323
Db			
Qy	1265	GTGAGCTCAAGCCAGGCGCCACAGCTGCACAGCCTGSCAAAGACCTGGGAGCAGCGGGC	1324
Db			
Qy	1324	TCCAGATCCCGGAGCGCCAGCCCAAACTCAGGACAAACGAGGGTGTCTGCTCACTGAG	1383
Db			
Qy	1325	TCCAGATCCCGGAGCGCCAGCCCAAACTCAGGACAAACGAGGGTGTCTGCTCACTGAG	1384
Db			
Qy	1384	AACTCAAGCCAGTGGATTTATGATACCGAGAAAGTCCACTGGGCGACGACCAAGCTC	1443
Db			
Qy	1385	AACTCAAGCCAGTGGATTTATGATACCGAGAAAGTCCACTGGGCGACGACCAAGCTC	1444
Db			
Qy	1444	CGCTGGGCGAGAGCTCCTTCGGAGAGTGCACAGGATGAGGACAAAGCAGCTGGCTTC	1503
Db			
Qy	1445	CGCTGGGCGAGAGCTCCTTCGGAGAGTGCACAGGATGAGGACAAAGCAGCTGGCTTC	1504
Db			
Qy	1504	CAGTGCCTCTCAAAAAGGTGCGCTGGAAAGTATTTCCGGCAGAGAGCTGATGGCATGT	1563
Db			
Qy	1505	CAGTGCCTCTCAAAAAGGTGCGCTGGAAAGTATTTCCGGCAGAGAGCTGATGGCATGT	1564
Db			
Qy	1564	GCAGGATGACCTCAACCCAGAAATTTGCTCCCTTTGATGAGAGCTGTGAGAAAGGCTTGG	1623
Db			
Qy	1565	GCAGGATGACCTCAACCCAGAAATTTGCTCCCTTTGATGAGAGCTGTGAGAAAGGCTTGG	1624
Db			
Qy	1624	GTCAACATCTTCATGAGAGTGTGGAGGTGGCTCCCTGGGCGAGCTGGTCAAGAGAGCAG	1683
Db			
Qy	1625	GTCAACATCTTCATGAGAGTGTGGAGGTGGCTCCCTGGGCGAGCTGGTCAAGAGAGCAG	1684
Db			
Qy	1684	GGCTGTCTCCAGAGGACCGGGCCCTGTATCTACCTGGGCGAGGCTGGAGGGTCTGGAA	1743
Db			
Qy	1685	GGCTGTCTCCAGAGGACCGGGCCCTGTATCTACTGGGCGAGGCTGGAGGGTCTGGAA	1744
Db			
Qy	1744	TACTCCACTCAAGAGATTTGTCATGGGAGCTCAAGCTGACAAAGCTGCTCTGTGCC	1803
Db			
Qy	1745	TACTCCACTCAAGAGATTTGTCATGGGAGCTCAAGCTGACAAAGCTGCTCTGTGCC	1804
Db			
Qy	1804	AGCGATGGGAGCCACGAGCCCTCTGTGATCTTTGGCCATGCTGTGCTCTCAACCTGAT	1863
Db			
Qy	1805	AGCGATGGGAGCCACGAGCCCTCTGTGATCTTTGGCCATGCTGTGCTCTCAACCTGAT	1864
Db			
Qy	1864	GGCTGGGAAAGTCTTTGCTCAAGGGGACTACATCTCTGGCAACAGAGACCAACATGGCT	1923
Db			

Db 1865 |||||GGCTGGGAAAGTCTTGTCTACAGGGGACTACATCCCTGGCCACAGAGACCCACATGGCT1924
Qy 1924 CCGGAGGTGGTGTCTGGCAGAGCTGCGACCGCAAGGTGGATGTCTGGCAGAGCTGCTGT1983
Db 1925 CCGGAGGTGGTGTCTGGCAGAGCTGCGACCGCAAGGTGGATGTCTGGCAGAGCTGCTGT1984
Qy 1984 ATGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAAGTCTCTCCGAGGGCCGCTC2043
Db 1985 ATGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAAGTCTCTCCGAGGGCCGCTC2044
Qy 2044 TGCCTCAAGATTGCCAGGAGCCTCCGCTGTGAGGGAGTCCCAACCTCTCTGGCCCT2103
Db 2045 TGCCTCAAGATTGCCAGGAGCCTCCGCTGTGAGGGAGTCCCAACCTCTCTGGCCCT2104
Qy 2104 CTCACAGCCAGGCCATCCAGAGGGGTGAGGAAGGCCATCCAGCGGTGTCTGCA2163
Db 2105 CTCACAGCCAGGCCATCCAGAGGGGTGAGGAAGGCCATCCAGCGGTGTCTGCA2164
Qy 2164 GCGGAGCTGGGAGGAAGTGAACCGGCACCTACAGCAAGTGGGAGGTCTGAAGAGCCCT2223
Db 2165 GCGGAGCTGGGAGGAAGTGAACCGGCACCTACAGCAAGTGGGAGGTCTGAAGAGCCCT2224
Qy 2224 TGGAGGGAGATATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATPACCAACAG2283
Db 2225 TGGAGGGAGATATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATPACCAACAG2284
Qy 2284 ACCCTCCATGCCAGCGAGAGCTTTGCCCAAGGGCCCCAGGGCCCCGGCCAGCTGAG2343
Db 2285 ACCCTCCATGCCAGCGAGAGCTTTGCCCAAGGGCCCCAGGGCCCCGGCCAGCTGAG2344
Qy 2344 GAGACAAAGGAGAGCCCTTAAGTCCAGCCTCTCTCCACAGAGCCGCCAGAGCCA2403
Db 2345 GAGACAAAGGAGAGCCCTTAAGTCCAGCCTCTCTCCACAGAGCCGCCAGAGCCA2404
Qy 2404 AACAAAGTCTCTCCCTTGAATTGTAGCAAGAGGAGTCTGGGATGTGGGAACCTTACCT2463
Db 2405 AACAAAGTCTCTCCCTTGAATTGTAGCAAGAGGAGTCTGGGATGTGGGAACCTTACCT2464
Qy 2464 CTGTCTCTCGGAGCAGCCCTTGCAGAAACCCAGTCCACAGAGGGGAAGCAACC2523
Db 2465 CTGTCTCTCGGAGCAGCCCTTGCAGAAACCCAGTCCACAGAGGGGAAGCAACC2524
Qy 2524 GTCCCGGAGCAGAACTGCGAGAGCTGGAAATAGAAATTATCTCAAGAGCCTGTCCAG2583
Db 2525 GTCCCGGAGCAGAACTGCGAGAGCTGGAAATAGAAATTATCTCAAGAGCCTGTCCAG2584
Qy 2584 CCATTTTCTCTGGAGAGCAGGACAAATTTCTCTGTGCTCAGCATCGACAGCCTCTCC2643
Db 2585 CCATTTTCTCTGGAGAGCAGGACAAATTTCTCTGTGCTCAGCATCGACAGCCTCTCC2644
Qy 2644 CTGTGGATGACAGTGAGAAAGCCATCAAGGCTCTCAAGCTCGCGGACACCCCTG2703
Db 2645 CTGTGGATGACAGTGAGAAAGCCATCAAGGCTCTCAAGCTCGCGGACACCCCTG2704
Qy 2704 AGCTAGGGGTACATCTCTGGAGCAGCCAGGCTCGAGGCTCGAAGCTTCCAGTCTGGAACATG2763
Db 2705 AGCTAGGGGTACATCTCTGGAGCAGCCAGGCTCGAAGCTTCCAGTCTGGAACATG2764
Qy 2764 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAGTCCAA2823
Db 2765 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAGTCCAA2824
Qy 2824 ATACAGTCTTTAAATGGTGAACCTGACATCCGGGAGTTCCACCGGTCAAGTGGGA2883
Db 2825 ATACAGTCTTTAAATGGTGAACCTGACATCCGGGAGTTCCACCGGTCAAGTGGGA2884
Qy 2884 GACATCGCAGCTGGCATCAGAGCCAGATCCAGCTGAGCCTTCAAGTCTGGTCAACAA2943
Db 2885 GACATCGCAGCTGGCATCAGAGCCAGATCCAGCTGAGCCTTCAAGTCTGGTCAACAA2944
Qy 2944 GACGGCAGCCTGTTCGCTACGACATGGAGGTGCCAGATCTGGGCAATCGACCTGCAAGTGC3003

Db 2945 GACGGCAGCCTGTTGCTACGACATGAGGTGCCAGACTCGGGATCGACTCGAGTGC3004
Qy 3004 ACATCGGCCCTGATGGCAGCTTCGCTGGAGCTTGGAGGTCAAAGCATGGCCAGCTGGAG3063
Db 3005 ACATCGGCCCTGATGGCAGCTTCGCTGGAGCTTGGAGGTCAAAGCATGGCCAGCTGGAG3064
Qy 3064 AACAGGCCCTAACCTGCTCCACCGCGGCTCCACACTCGCGGAAACAGCCTTCCTG3123
Db 3065 AACAGGCCCTAACCTGCTCCACCGCGGCTCCACACTCGCGGAAACAGCCTTCCTG3124
Qy 3124 CTCGCTGACCATGCTGCTTGAACACAGAGCTCAGCGTTCCAGGGGATTCGCAGCC3183
Db 3125 CTCGCTGACCATGCTGCTTGAACACAGAGCTCAGCGTTCCAGGGGATTCGCAGCC3184
Qy 3184 CCCCGCTTCAAGTGGGAAACAGGCGCTTCGAGAGCAAGGTGGGGCAAGCAAGTGC3243
Db 3185 CCCCGCTTCAAGTGGGAAACAGGCGCTTCGAGAGCAAGGTGGGGCAAGCAAGTGC3244
Qy 3244 TCCAGGATTTTCAACCTGAGCCCTGCCCCACCTGTGAAAAAACAATCCGCCAGGTGAA3303
Db 3245 TCCAGGATTTTCAACCTGAGCCCTGCCCCACCTGTGAAAAAACAATCCGCCAGGTGAA3304
Qy 3304 GAGACAGAGAGATGCGCAGAGTTACTTGGGNAACAAACAGGGATCTTTTCTGCC3363
Db 3305 GAGACAGAGAGATGCGCAGAGTTACTTGGGNAACAAACAGGGATCTTTTCTGCC3364
Qy 3364 CCTGCTCAGTTCGAGTTGGCTGAGCCCGTGGATTCAGTGACCAATTTGTGGCAGACAGG3423
Db 3365 CCTGCTCAGTTCGAGTTGGCTGAGCCCGTGGATTCAGTGACCAATTTGTGGCAGACAGG3424
Qy 3424 GGAGAGCAGCTTCCAGCCTGGGTGAGAGGGTGGGGAGCCCTTGGCCCTCTCACCTC3483
Db 3425 GGAGAGCAGCTTCCAGCCTGGGTGAGAGGGTGGGGAGCCCTTGGCCCTCTCACCTC3484
Qy 3484 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGCCCAAACTCAGGTTCAAGTGCAGAACCA3543
Db 3485 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGCCCAAACTCAGGTTCAAGTGCAGAACCA3544
Qy 3544 GGTCAAGAGTATGCCCGCTAGTTTAAAGGGGCCCTCTAAACCCCTTGGCTGGCCTC3603
Db 3545 GGTCAAGAGTATGCCCGCTAGTTTAAAGGGGCCCTCTAAACCCCTTGGCTGGCCTC3604
Qy 3604 ACCTGGCAGCTACCCCTTTTGGGTGAGGGGAAAAAGATGCTGAGACCTTGGGAGGCT3663
Db 3605 ACCTGGCAGCTACCCCTTTTGGGTGAGGGGAAAAAGATGCTGAGACCTTGGGAGGCT3664
Qy 3664 CCCTGGTAGAATACACACATTTTTCAGGTTGTGCAACAGAGTCTTGAAGTTCAGCTCT3723
Db 3665 CCCTGGTAGAATACACACATTTTTCAGGTTGTGCAACAGAGTCTTGAAGTTCAGCTCT3724
Qy 3724 GGTTCAGCAAGGACCAAGAGGTGTGTAAAGTGAAGTGGTTCTCAGTCCCCAGACATGT3783
Db 3725 GGTTCAGCAAGGACCAAGAGGTGTGTAAAGTGAAGTGGTTCTCAGTCCCCAGACATGT3784
Qy 3784 GCGCCCTTTGCTGGCTACCACTTTCAGAGCAGAGGCCCCCGAGCCCTTTCAGGC3843
Db 3785 GCGCCCTTTGCTGGCTACCACTTTCAGAGCAGAGGCCCCCGAGCCCTTTCAGGC3844
Qy 3844 CCAGCATGCCCCAGACTCGCTGGCACTCAGTTCCTCATCTGTAAAGTGAAGGTGAT3903
Db 3845 CCAGCATGCCCCAGACTCGCTGGCACTCAGTTCCTCATCTGTAAAGTGAAGGTGAT3904
Qy 3904 CGAGGATATGCTCAGAGAACAGTCTGTGGATGGACATGATCAGTGTCCAGGAAAGCAG3963
Db 3905 CGAGGATATGCTCAGAGAACAGTCTGTGGATGGACATGATCAGTGTCCAGGAAAGCAG3964
Qy 3964 CAGAGAGAGCCTCCGGGCCCCAGCCCACTATCAGTGTCCAGGCTGTGTTCCCGAG4023
Db 3965 CAGAGAGAGCCTCCGGGCCCCAGCCCACTATCAGTGTCCAGGCTGTGTTCCCGAG4024
Qy 4024 AGCAGAGCTCAGCATCACTGACACTCACCTGCTGCCCTGCCCCAGAGGGTACTGC4083
Db 4025 AGCAGAGCTCAGCATCACTGACACTCACCTGCTGCCCTGCCCCAGAGGGTACTGC4084

Db	3355		TGCCCTGCTCCAGTCGAGTTGGCCCTGACCCCTTGGATTCAGTGACCATTTGCTGGCAGAC	3414
Qy	3421		AGGGGAGAGCAGCTTCCAGCCTCGGTGAGAAAGGGGTGGCGAGGCCCTTGGCCCCCTCACC	3480
Db	3415		AGGGGAGAGCAGCTTCCAGCCTCGGTGAGAAAGGGGTGGCGAGGCCCTCGCCCCCTCACC	3474
Qy	3481		CTCAGAGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCAAATCTCAGGTTTCAGTGCAGAA	3540
Db	3475		CTCAGAGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCAAATCTCAGGTTTCAGTGCAGAA	3534
Qy	3541		CCAGGTTCAGCAGGTATGAGCCGCCGTAGTTTAAAGGGGCCCTTAAACCCCTTGCCTGGC	3600
Db	3535		CCAGGTTCAGCAGGTATGAGCCGCCGTAGTTTAAAGGGGCCCTTAAACCCCTTGCCTGGC	3594
Qy	3601		CTCACCTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAGAAATGCTTGACCTCGGGAAG	3660
Db	3595		CTCACCTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAGAAATGCTTGACCTCGGGAAG	3654
Qy	3661		GCTCCCTGGTAGAATACACACACTTTTCAGGTTGTGCAACACAGGTCTGTAGTTGACC	3720
Db	3655		GCTCCCTGGTAGAATACACACACTTTTCAGGTTGTGCAACACAGGTCTGTAGTTGACC	3714
Qy	3721		TCGTGGTTTCAGCCAGGACCAAGAGGTGTCTAAGTGAAGTGGTTCAGTCTCCACACA	3780
Db	3715		TCGTGGTTTCAGCCAGGACCAAGAGGTGTCTAAGTGAAGTGGTTCAGTCTCCACACA	3774
Qy	3781		TGTGCCCTTTGCTGCTGGCTTACCACCTCTTCCCCAGAGCAGCAGGCCCCGAGCCCTTCA	3840
Db	3775		TGTGCCCTTTGCTGCTGGCTTACCACCTCTTCCCCAGAGCAGCAGGCCCCGAGCCCTTCA	3834
Qy	3841		GGCCCCAGCATGCCCCAGACTCGTGGCACTCAGTTCCTCATCTGTAAAGGTGAAGGGT	3900
Db	3835		GGCCCCAGCATGCCCCAGACTCGTGGCACTCAGTTCCTCATCTGTAAAGGTGAAGGGT	3894
Qy	3901		GATGCAGGATATGCTTCACAGGAACACTGTGTGATGGAATGATGATCAGTCTGAAGAAAG	3960
Db	3895		GATGCAGGATATGCTTCACAGGAACACTGTGTGATGGAATGATGATCAGTCTGAAGAAAG	3954
Qy	3961		CAGCAGAGAGAGACG-TCCGGCGGCCCCAGCCCCCACTATCAGTGTCCAGCGTGTGTTTC	4019
Db	3955		CAGCAGAGAGAGACGCTCCGGCGGCCCCAGCCCCCACTATCAGTGTCCAGCGTGTGTTTC	4014
Qy	4020		CCAGAGCACAGCTCAGCATCACACTGACACTCACCTCGCCCTGCCCCCTGGCCAGAGGGTA	4079
Db	4015		CCAGAGCACAGCTCAGCATCACACTGACACTCACCTCGCCCTGCCCCCTGGCCAGAGGGTA	4074
Qy	4080		CTGCCGAGGCACTTTGCATCTGTATGACCTCAAAGCACTTTATGCTGCCCTCTGGCA	4139
Db	4075		CTGCCGAGGCACTTTGCATCTGTATGACCTCAAAGCACTTTATGCTGCCCTCTGGCA	4134
Qy	4140		GGGCAGGGCAGGGCAGTGCACACTGTAGAGCATAGCAAGCAGCAGATGGGTGAAGGGA	4199
Db	4135		GGGCAGGGCAGGGCAGTGCACACTGTAGAGCATAGCAAGCAGCAGATGGGTGAAGGGA	4194
Qy	4200		CACAGCTTTGAGCTGTCCACATGATGTGACTCCTCAAACCTCTTCCAGATTTCTCTAAG	4259
Db	4195		CACAGCTTTGAGCTGTCCACATGATGTGACTCCTCAAACCTCTTCCAGATTTCTCTAAG	4254
Qy	4260		AATAGCACCCCTTTCCCAATTGCCCAAGCTTAGCTCTTCTCCAGGGGAGCTACTCAGG	4319
Db	4255		AATAGCACCCCTTTCCCAATTGCCCAAGCTTAGCTCTTCTCCAGGGGAGCTACTCAGG	4314
Qy	4320		ACTCACGTAGCATTAATCAGCTGTGAATCGTTCAGGGGGTGTCTGTAGCCTCAACCTCC	4379
Db	4315		ACTCACGTAGCATTAATCAGCTGTGAATCGTTCAGGGGGTGTCTGTAGCCTCAACCTCC	4374
Qy	4380		TGGGGCAGGGGACGCCGAGACTCCGTGGGAGAGCTCATTTCCACATCTTTGCCAAGACAG	4439
Db	4375		TGGGGCAGGGGACGCCGAGACTCCATGGGAGAGCTCATTTCCACATCTTTGCCAAGACAG	4434
Qy	4440		CCTTTGTTCAGCTGTCCACATTCAGTTCAGACTGCTCCCGGGGAGAGAGCCCGCCCCCA	4499

Db	4435	CGTTTGTCCAGCTGTCCACATTTGAGTCAGACTCTCTCCCGGGGAGAGAGCGCCCGCCCCCA	4499		
Qy	4500	GCACATAAAGAACTCGAGCCCTTGGTACTCGACAGTCTGGGTTGTAGAGAACTCTTTGTAA	4559		
Db	4495	GCACATAAAGAACTCGAGCCCTTGGTACTCGACAGTCTGGGTTGTAGAGAACTCTTTGTAA	4554		
Qy	4560	GCATAAAGTTTGGGGTGATGACAAATGTT	4589		
Db	4555	GCATAAAGTTTGGGGTGATGACAAATGTT	4584		
RESULT 6	BC035576	4452 bp	mrna	linear	PRI 06-OCT-2003
LOCUS	BC035576				
DEFINITION	Homo sapiens mitogen-activated protein kinase kinase 14, mRNA (CDNA clone MGC:45335 IMAGE:5497185), complete cds.				
ACCESSION	BC035576	1	GI:23272579		
VERSION	BC035576				
KEYWORDS	MGC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 4452)				
	Srausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,K.H., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zebberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaby,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Musny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzyzanski,M.I., Skalska,U., Smalley,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
MEDLINE	22388257				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 4452)				
AUTHORS	Srausberg,R.				
TITLE	Direct Submission				
JOURNAL	Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgabbs-r@mail.nih.gov Tissue Procurement: Lou Staudt cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nigri.nih.gov Akhter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Mastello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Scantripoop,S., Thomas,P.J., Touchman,J.W., Tsurgou,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A.C., Zhang,L.-H. and Green,E.D.				

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 69 Row: f Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505396.

FEATURES

source

Location/Qualifiers
1..4452
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:45335 IMAGE:5497185"
/issue_type="Lymph, lymphoma"
/clone_lib="NIH MGC_85"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

gene

1..4452
/gene="MAP3K14"
/note="synonyms: NIK, HSN1K, HS, FTDCL1B"
/db_xref="LocusID:9020"
/db_xref="MIM:604655"
71..2914
/gene="MAP3K14"

CDS

1..2914
/product="MAP3K14 protein"
/protein_id="AAH35576.1"
/db_xref="GI:23272580"
/db_xref="LocusID:9020"
/db_xref="MIM:604655"
/translation="MAVMEACPGAGSVAQQKELPKAKETPLGLKQSSVYKLEA
VEKSPVFGKWLNDVITKGTAKGSEBAGPAAISIIQAECENSOEFSPTFSRIFI
AGSKOYSOSIEDQIPNNVAHATEGMARVCKWKRRSKARKKRRKSKSLAHAGVA
LAKPLPRPEQESCTIPVEDSPILGAPVVRNTPOFTKPLKPGQLCFKQLGRLR
PALPSELHLISPLQCLNHVWKLHPDQGGPLPFTHPFYSRLPHFPPLQPPWK
PHLPSEFGLKACVDSQPLPDPLHSLKACVDSQPLPGLPHLEPCLSRGAHEKFSVE
EYLVAHQSSVQGAHSLTSLAKTWAARGSRSEPSKPTEDNEGVLLTEKLKPYDYE
YREYVHAQTHOIRLQSGSGFGEVHRMEDKQTFQCAVKKVRLEVFRAEELMACGLTSP
RIVPLGYAVREGPWNIEMELLEGLGVLQVKEQCLPEDRALYVLGOALEGLVLS
RIILHGDVADNVLLSSDGHAAALCDFGHVNCLOPDGLGKSLITGDYIRGETHWAP
VVLGRSDAKVDWSSCMHLMLNGCHPWTQFFRGLCLKIASFPVPREIIPSCAP
LTAQAIQAGRELKREP IHRVSAEELGGKVRNALQQVGLKSPWRGEKPEPRHPPNQANY
HOTLAPRELPRAPRPRAETTGAPLQPLPPLPPFPKPSPPITLSEESGMW
EPLLSLEPAPRAPNSPERKATPEOBLQOLELFLNLSQPSFSLBOQLISCL
STDLSLSDSKSPKASQSSRDVTLSSVHSSWSQASRSSNNMVLARGPTDTPS
YNGVKVQIQSLNGELHLIRFHRVKVGDIGTGISQIPAAAFSLVTKDGPVPRDME
VPDGGIDLOCTUAPDGSFAWSRVRXHGQENRP"
1286..2011
/gene="MAP3K14"
/note="S_TKc; Region: Serine/Threonine protein kinases,
catalytic domain"
/db_xref="CDD:smart00220"

misc_feature

92.5%; Score 4249.4; DB 9; Length 4452;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 4354; Conservative 0; Mismatches 41; Indels 10; Gaps 7;
Query Match

ORIGIN

Query Match 92.5%; Score 4249.4; DB 9; Length 4452;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 4354; Conservative 0; Mismatches 41; Indels 10; Gaps 7;
Query 194 GTGTTTGGTCTCTCTCAGGATGAGCACAAAGCCTGGGAGATGGCAGTGTGGAATGGCCT 253
Db 33 GAGATCGGCGCTTGGCGCGATGAGCACAAAGCCTGGGAGATGGCAGTGTGGAATGGCCT 92
Query 254 GCCCAGGTCCCTGGCTCAGCAGTGGGCGACGAGGAACCTCCCAAGCCAAAGGAGA 313
Db 93 GCCCAGGTCCCTGGCTCAGCAGTGGGCGACGAGGAACCTCCCAAGCCAAAGGAGA 152
Query 314 AGACGCGCGCACTGGGGAAGAAACAGAGCTCCGTCTACAAAGCTTGAGGCGGTGGAGAGA 373
Db 153 AGACGCGCGCACTGGGGAAGAAACAGAGCTCCGTCTACAAAGCTTGAGGCGGTGGAGAGA 212
Query 374 GCCCTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGACCTGATTTACCAAGGGGCACAGCCA 433
Db 213 GCCCTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGACCTGATTTACCAAGGGGCACAGCCA 272

434 AGGAAGGCTCCGAGGCGAGGCGCAGTGCATCTCTATCATCGCCAGGCTGAGTGTGAGA 493
Db 273 AGGAAGGCTCCGAGGCGAGGCGCAGTGCATCTCTATCATCGCCAGGCTGAGTGTGAGA 332
Query 494 ATAGCCAAAGTTCAGGCCACCTTTTTCAGAACGATTTTTCATCGCTGGTCCCAACAGT 553
Db 333 ATAGCCAAAGTTCAGGCCACCTTTTTCAGAACGATTTTTCATCGCTGGTCCCAACAGT 392
Query 554 ACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATCTCTACAGAGGCA 613
Db 393 ACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATCTCTACAGAGGCA 452
Query 614 AAATGGCCCGTGTGTGTTGGAAAGGAAAGCTGCGAGCAAAAGCCCGGAAAGAAACGGAAGA 673
Db 453 AAATGGCCCGTGTGTGTTGGAAAGGAAAGCTGCGAGCAAAAGCCCGGAAAGAAACGGAAGA 512
Query 674 AGAAGAGTCAAAAGTCCCTGGCTCATGCGAGAGTGGCTTGGCCAAAACCCCTCCCAAGA 733
Db 513 AGAAGAGTCAAAAGTCCCTGGCTCATGCGAGAGTGGCTTGGCCAAAACCCCTCCCAAGA 572
Query 734 CCCCTGAGCAGGAGAGTGCACCATCCCAAGTGCAGAGGATGAGTCTCCACTTCGGCGGCC 793
Db 573 CCCCTGAGCAGGAGAGTGCACCATCCCAAGTGCAGAGGATGAGTCTCCACTTCGGCGGCC 632
Query 794 CATATGTTAGAAACACCCCGCAGTTTCAACAAAGCCTCTGAAGGAAACAGGCCCTTGGGCAAC 853
Db 633 CATATGTTAGAAACACCCCGCAGTTTCAACAAAGCCTCTGAAGGAAACAGGCCCTTGGGCAAC 692
Query 854 TCTGTTTAAAGCAGTTCGCGAGGGCTTACGGCCGGCTCTGGCTCGATCAGAACTCCACA 913
Db 693 TCTGTTTAAAGCAGTTCGCGAGGGCTTACGGCCGGCTCTGGCTCGATCAGAACTCCACA 752
Query 914 AACTGATCAGCCCTTGAATGTCTGAACACAGTGTGGAACCTGACCAACCCCAAGGACG 973
Db 753 AACTGATCAGCCCTTGAATGTCTGAACACAGTGTGGAACCTGACCAACCCCAAGGACG 812
Query 974 GAGGCCCCCTGCCCTGCGCCACGACCCCTTCCCCCTATAGCAGACTGCTCATCCCTTCC 1033
Db 813 GAGGCCCCCTGCCCTGCGCCACGACCCCTTCCCCCTATAGCAGACTGCTCATCCCTTCC 872
Query 1034 CATTCCACCTCTCAGCCCTGGAAACCTCACCTCTGAGTCTCTTCTTGGCAAACTGG 1093
Db 873 CATTCCACCTCTCAGCCCTGGAAACCTCACCTCTGAGTCTCTTCTTGGCAAACTGG 932
Query 1094 CCTGTGTAGACAGCCAGAAAACCTTGTGCTGACCCACACTGAGCAAACTGGGCTGTGTAG 1153
Db 933 CCTGTGTAGACAGCCAGAAAACCTTGTGCTGACCCACACTGAGCAAACTGGGCTGTGTAG 992
Query 1154 ACAGTCCAAAAGCCCTGCTGCGCCACACACTGGAGCCAGCTGCTGTCTGTGTGCCC 1213
Db 993 ACAGTCCAAAAGCCCTGCTGCGCCACACACTGGAGCCAGCTGCTGTCTGTGTGCCC 1052
Query 1214 ATGAGAAATTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGCGAGCGTGAATCA 1273
Db 1053 ATGAGAAATTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGCGAGCGTGAATCA 1112
Query 1274 GCCAGGCCACAGCTGCTGCGCCAGACCTGGGCGACGCGGCTCCAGATCCC 1333
Db 1113 GCCAGGCCACAGCTGCTGCGCCAGACCTGGGCGACGCGGCTCCAGATCCC 1172
Query 1334 GGGAGCCCGCCCAAACTGAGGACAAACGAGGGTGTCTGTCTCACTGAGAACTCAAGC 1393
Db 1173 GGGAGCCCGCCCAAACTGAGGACAAACGAGGGTGTCTGTCTCACTGAGAACTCAAGC 1232
Query 1394 CAGTGGATTTATGAGTACCGAGAAAGTCCACTGCGGCCACGCAAGCTCCGCTGGGCA 1453
Db 1233 CAGTGGATTTATGAGTACCGAGAAAGTCCACTGCGGCCACGCAAGCTCCGCTGGGCA 1292
Query 1454 GAGGCTCCTTTCGAGAGGTTGCAAGGATGGAGGACGAGCAACTGGCTTCCAGTGGCTG 1513
Db 1293 GAGGCTCCTTTCGAGAGGTTGCAAGGATGGAGGACGAGCAACTGGCTTCCAGTGGCTG 1352
Query 1514 TCAAAAAGGTGCGCCTGGAAAGTATTTTCGGGCGACGAGGAGCTGATGGCATGTGAGGATTGA 1573

Db 1353 TCAAAAAGTGGCTGGAAGTATTTGGGAGAGGAGCTGATGGCATGTGCAGATTGA 1412
Qy 1574 CCTCACCAGAAATTGTCTCCCTTTGTATGAGAGTGTGAGAGAGGCTTTGGGTCAACATCT 1633
Db 1413 CCTCACCAGAAATTGTCTCCCTTTGTATGAGAGTGTGAGAGAGGCTTTGGGTCAACATCT 1472
Qy 1634 TCATGGAGCTCTGGAAGGTGGCTCCCTGGGCCAGCTGTGTCAAGAGAGAGGCTGTCTCC 1693
Db 1473 TCATGGAGCTCTGGAAGGTGGCTCCCTGGGCCAGCTGTGTCAAGAGAGAGGCTGTCTCC 1532
Qy 1694 CAGAGGACCGGGCCCTGTACTCTGTGGCCAGGSCCTGGAGGGTCTGGAATACCTCCACT 1753
Db 1533 CAGAGGACCGGGCCCTGTACTCTGTGGCCAGGSCCTGGAGGGTCTGGAATACCTCCACT 1592
Qy 1754 CACGAAGGATTTCTGATGGGACGCTCAAGCTGACAACTGCTCTCTGTCCAGCGATGGGA 1813
Db 1593 CACGAAGGATTTCTGATGGGACGCTCAAGCTGACAACTGCTCTCTGTCCAGCGATGGGA 1652
Qy 1814 GCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTCTTTCAACCTGATGGCTGGGAA 1873
Db 1653 GCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTCTTTCAACCTGATGGCTGGGAA 1712
Qy 1874 AGTCTTGTCTCAGAGGGAGTACATCCCTGSCACAGAGACCCACATGSCCTCCGAGGTGG 1933
Db 1713 AGTCTTGTCTCAGAGGGAGTACATCCCTGSCACAGAGACCCACATGSCCTCCGAGGTGG 1772
Qy 1934 TGCTGGGAGGAGCTGCGACGCCCAAGGTGGATGTCTGGAGCAGCTGCTGTATGATGTGC 1993
Db 1773 TGCTGGGAGGAGCTGCGACGCCCAAGGTGGATGTCTGGAGCAGCTGCTGTATGATGTGC 1832
Qy 1994 ACATGCTCAACGGGTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTCTGCTCAAGA 2053
Db 1833 ACATGCTCAACGGGTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTCTGCTCAAGA 1892
Qy 2054 TTGCGAGGAGCCTCCGCTGTGAGGAGATCCACCCCTCTGCGCCCTCTCAAGGCC 2113
Db 1893 TTGCGAGGAGCCTCCGCTGTGAGGAGATCCACCCCTCTGCGCCCTCTCAAGGCC 1952
Qy 2114 AGGCGATCAAGAGGGGTGAGGAAGAGGCCATCAACCGGTCTCTGCGAGCGAGTGG 2173
Db 1953 AGGCGATCAAGAGGGGTGAGGAAGAGGCCATCAACCGGTCTCTGCGAGCGAGTGG 2012
Qy 2174 GAGGAAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGAGCCCTTTGGAGGGAG 2233
Db 2013 GAGGAAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGAGCCCTTTGGAGGGAG 2072
Qy 2234 AATATAAGAAACCAAGACATCCACCGCAATCAAGCCAAATTACCAAGACCCCTCCATG 2293
Db 2073 AATATAAGAAACCAAGACATCCACCGCAATCAAGCCAAATTACCAAGACCCCTCCATG 2132
Qy 2294 CCCAGCGAGAGCTTTCCGCAAGGGCCCGAGGGCCCGCCAGCTGAGGAGACAAG 2353
Db 2133 CCCAGCGAGAGCTTTCCGCAAGGGCCCGAGGGCCCGCCAGCTGAGGAGACAAG 2192
Qy 2354 GCAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAGAGCAAAAGTCTC 2413
Db 2193 GCAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAGAGCAAAAGTCTC 2252
Qy 2414 CTCCCTTGTACTTTGAGCAAGGAGGTCTGGGATGTGGGAACCCCTTACCTCTGTCCTCC 2473
Db 2253 CTCCCTTGTACTTTGAGCAAGGAGGTCTGGGATGTGGGAACCCCTTACCTCTGTCCTCC 2312
Qy 2474 TGGAGCAGCCCTTCCAGAAACCCAGCTCACAGAGCGGAAGCAACCGTCCCGAGC 2533
Db 2313 TGGAGCAGCCCTTCCAGAAACCCAGCTCACAGAGCGGAAGCAACCGTCCCGAGC 2372
Qy 2534 AGGAACGTGACAGCTGGAATAGAAATTATTCCTCAACAGGCTCTGCCAGCAATTTTCTC 2593
Db 2373 AGGAACGTGACAGCTGGAATAGAAATTATTCCTCAACAGGCTCTGCCAGCAATTTTCTC 2432
Qy 2594 TGGAGGAGCAGGAGCAAAATTTCTCTGTCCTCAGCATCGACAGCTCTCCCTGTCCGATG 2653

Db 2433 TGGAGGAGCAGGAGCAAAATTTCTCTGTGCTCTCAGCATCGACAGCTCTCCCTGTCCGATG 2492
Qy 2654 ACAGTGAAGAAAGACCCATCAAAAGCCTCTCAAAAGCTCGCGGGACACCTCTAGCTCAGGCG 2713
Db 2493 ACAGTGAAGAAAGACCCATCAAAAGCCTCTCAAAAGCTCGCGGGACACCTCTAGCTCAGGCG 2552
Qy 2714 TACACTCTGAGCAGCAGCCGAGGCTGGAAGCTCAGCTGGAACATGCTGTCTGGCCC 2773
Db 2553 TACACTCTGAGCAGCAGCCGAGGCTGGAAGCTCAGCTGGAACATGCTGTCTGGCCC 2612
Qy 2774 GGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAAGTCCAAATACAGTCTC 2833
Db 2613 GGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAAGTCCAAATACAGTCTC 2672
Qy 2834 TTAATGTGTAACACTCTGCACATCCGGAGTTTCCACCGGCTCAAAAGTGGGAGACATCGCA 2893
Db 2673 TTAATGTGTAACACTCTGCACATCCGGAGTTTCCACCGGCTCAAAAGTGGGAGACATCGCA 2732
Qy 2894 CTGSCATCAGCAGCAGATCCAGCTCAGCTTTCAGCTTGGTCAACAAAGACGGGACG 2953
Db 2733 CTGSCATCAGCAGCAGATCCAGCTCAGCTTTCAGCTTGGTCAACAAAGACGGGACG 2792
Qy 2954 CTGTTGCTACGACATGGAGTTCAGACTCGGCGATCGACCTGACATGACATGACCTGGCCC 3013
Db 2793 CTGTTGCTACGACATGGAGTTCAGACTCGGCGATCGACCTGACATGACATGACCTGGCCC 2852
Qy 3014 CTGATGSCAGCTTCGCTGAGCTGGAGGGTCAAGCATGCGCAGCTGAGAACAGGCGCT 3073
Db 2853 CTGATGSCAGCTTCGCTGAGCTGGAGGGTCAAGCATGCGCAGCTGAGAACAGGCGCT 2912
Qy 3074 AACCTGCTTCCACCCCGCTCCACACTGCGGAAAGCAGCTTCTGCTCGGTGAC 3133
Db 2913 AACCTGCTTCCACCCCGCTCCACACTGCGGAAAGCAGCTTCTGCTCGGTGAC 2971
Qy 3134 GATGCTGCTGAAACACAGGCTCAGCCGTTCCAGGSGATTCGCGACCCCGGCTCA 3193
Db 2972 GATGCTGCTGAAACACAGGCTCAGCCGTTCCAGGSGATTCGCGAG - CCCCCTGGCTCA 3029
Qy 3194 CAGTGGGAAACAGGGCTTCGAGCAGCAAGGTGGGGCAAGCAGAAATGCTCCAGGAT 3253
Db 3030 GCAGTGGGAAACAGGGCTTCGAGCAGCAAGGTGGGGCAAGCAGAAATGCTCCAGGAT 3089
Qy 3254 TCACACTGAGCTGCTGCCACCTGCTGTAAGAAACA - TCGGCCACTGTAAGACAGAA 3312
Db 3090 TCACACTGAGCTGCTGCCACCTGCTGTAAGAAACA - TCGGCCACTGTAAGACAGAA - A 3148
Qy 3313 GAGGATGAGCAGAGTTTACCTGGGAAACAAACAGGAGTCTTTTCTGCCCCCTGCTCCA 3372
Db 3149 GAGGATGAGCAGAGTTTACCTGGGAAACAAACAGG - ---ATCTTCTCTGCCCCCTGCTCCA 3204
Qy 3373 GTCGAGTTGGCTCAGCCGCTTGGATCAGTGACCATTTGTCGAGACAGGGGAGAGCAG 3432
Db 3205 GTCGAGTTGGCTCAGCCGCTTGGATCAGTGACCATTTGTCGAGACAGGGGAGAGCAG 3264
Qy 3433 CTTCCAGCTGGGTGAGAGGGGTGGCGGAGCCCTTCGGGCCCTCACCCCTCAGGCTGCT 3492
Db 3265 CTTCCAGCTGGGTGAGAGGGGTGGCGGAGCCCTTCGGGCCCTCACCCCTCAGGCTGCT 3324
Qy 3493 GTGAGAGTGTCAAGTGTAAAGGGCCCAACTCAGGTTTCAGTGAGAAACAGGTCAGGAG 3552
Db 3325 GTGAGAGTGTCAAGTGTAAAGGGCCCAACTCAGGTTTCAGTGAGAAACAGGTCAGGAG 3384
Qy 3553 GTATGCCCGCTGAGTTAAAGGGGCCCTTAAACCCCTTGCCTGCTCACCTGAGGCA 3612
Db 3385 GTATGCCCGCTGAGTTAAAGGGGCCCTTAAACCCCTTGCCTGCTCACCTGAGGCA 3444
Qy 3613 GCTCACCCCTTTGGGTGAGGGGAAAGAAATGCTGACCCCTGGGAAAGGCTCCCTGGTAG 3672
Db 3445 GCTCACCCCTTTGGGTGAGGGGAAAGAAATGCTGACCCCTGGGAAAGGCTCCCTGGTAG 3504
Qy 3673 AATACACCACTTTTTCAGGTTGTGCAACAGAGGTCCTGAGTTGACTCTGTGTTAGCC 3732
Db 3505 AATACACCACTTTTTCAGGTTGTGCAACAGAGTCTCTGAGTTGACTCTGTGTTAGCC 3564

```
QY 3733 AAGGACCAAGAGGTGTGAAGTGAAGTGTCTCAGTCCCGACACATGTGCCCTTTG 3792
Db 3565 AAGGACCAAGAGGTGTGAAGTGAAGTGTCTCAGTCCCGACACATGTGCCCTTTG 3624
QY 3793 CTGCTGGCTACCACTCTTCCCGACAGCAGGCCCCGAGCCCCCTTCCAGGCCCAAGCACTG 3852
Db 3625 CTGCTGGCTACCACTCTTCCCGACAGCAGGCCCCGAGCCCCCTTCCAGGCCCAAGCACTG 3684
QY 3853 CCCGACACTCGTGGCACTCAGTTCCTCCATCTGTAAGGTGAAGGTGATGCAGGATAT 3912
Db 3685 CCCGACACTCGTGGCACTCAGTTCCTCCATCTGTAAGGTGAAGGTGATGCAGGATAT 3744
QY 3913 GCCTCAGCAGGAACAGTCTGTGATGAGCATGATCAGTCTAAGGAAACAGCAGAGAGAG 3972
Db 3745 GCCTCAGCAGGAACAGTCTGTGATGAGCATGATCAGTCTAAGGAAACAGCAGAGAGAG 3804
QY 3973 ACG-TCCGGCGCCCGACGCCCACTATCAGTGTCCAGCGTGTGGTTCCCGACAGACAGC 4031
Db 3805 ACGCTCCGGCGCCCGACGCCCACTATCAGTGTCCAGCGTGTGGTTCCCGACAGACAGC 3864
QY 4032 TCAGCATCAGCATGACATCACCCTGCCCTGCCCTGCGCAGAGGGTACTGCCACGGCA 4091
Db 3865 TCAGCATCAGCATGACATCACCCTGCCCTGCCCTGCGCAGAGGGTACTGCCACGGCA 3924
QY 4092 CTTTGCACCTCTGATGACCTCAAAGCACATTTTCATGTGCTGCCCTCTGCCAGGGCAGGCAGG 4151
Db 3925 CTTTGCACCTCTGATGACCTCAAAGCACATTTTCATGTGCTGCCCTCTGCCAGGGCAGGCAGG 3984
QY 4152 GCAGTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGAACACAGTCTTGAG 4211
Db 3985 GCAGTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGAACACAGTCTTGAG 4044
QY 4212 CTGTCCACATGATGTGACTCTCAAACTCTTCAGATTTCTTAAGATAGCACCCCC 4271
Db 4045 CTGTCCACATGATGTGACTCTCAAACTCTTCAGATTTCTTAAGATAGCACCCCC 4104
QY 4272 TTCCCCATTGCCCGAGCTTAGCTCTTCTCCAGGGAGCTACTCAGGACTCACGTAGCA 4331
Db 4105 TTCCCCATTGCCCGAGCTTAGCTCTTCTCCAGGGAGCTACTCAGGACTCACGTAGCA 4164
QY 4332 TTAATCAGCTGTGAATCGTCAGGGGTGTCTGTAGCCTCAACCTCTCGGGGCGAGGGA 4391
Db 4165 TTAATCAGCTGTGAATCGTCAGGGGTGTCTGTAGCCTCAACCTCTCGGGGCGAGGGA 4224
QY 4392 CGCCGAGACTCGTGGGAGAAGCTCATTCCTCCATCTTCCGACAGAGCCCTTGTCCAGC 4451
Db 4225 CGCCGAGACTCGTGGGAGAAGCTCATTCCTCCATCTTCCGACAGAGCCCTTGTCCAGC 4284
QY 4452 TGTCCACATTTAGTCAAGTCTGCTCCGGGAGAGAGCCCGGCCCCCGACACATAAGAA 4511
Db 4285 TGTCCACATTTAGTCAAGTCTGCTCCGGGAGAGAGCCCGGCCCCCGACACATAAGAA 4344
QY 4512 CTGCAGCCTTGGTACTGACAGTCTGGGTGTAGAGAACTCTTTGTAGCAATAAAGTTT 4571
Db 4345 CTGCAGCCTTGGTACTGACAGTCTGGGTGTAGAGAACTCTTTGTAGCAATAAAGTTT 4404
QY 4572 GGGGTGATCAAAAAGTTTAAAAAAA 4596
Db 4405 GGGGTGATCAAAAAGTTTAAAAAAA 4429
```

```
RESULT 7
AR062290
LOCUS AR062290 3156 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5843721.
ACCESSION AR062290
VERSION AR062290.1 GI:5989981
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3156)
```

AUTHORS Rothe,M. and Wu,L.
TITLE Nucleic acids encoding human NIK protein
JOURNAL Patent: US 5843721-A 1 01-DEC-1998;
FEATURES Location/Qualifiers
source 1..3156
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

```
Query Match 68.5%; Score 3149.6; DB 6; Length 3156;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 232 ATGCGAGTGTATGGAATGGCTGCTGCCAGGTGCCCTCGCTCAGCATGTGGGCGACGAGAAG 291  
Db 1 ATGCGAGTGTATGGAATGGCTGCTGCCAGGTGCCCTCGCTCAGCATGTGGGCGACGAGAAG 60  
QY 292 GAACTCCCAAGCCAAAGGAGAGACGCCCGCATCTGGGGAAGAAACAGAGCTCCGTCTAC 351  
Db 61 GAACTCCCAAGCCAAAGGAGAGACGCCCGCATCTGGGGAAGAAACAGAGCTCCGTCTAC 120  
QY 352 AAGCTTGGGCGCGTGGAGAGAGAGCCCTGTGTTCTCGGAAAGTGGAGATCCTCAATGAC 411  
Db 121 AAGCTTGGGCGCGTGGAGAGAGAGCCCTGTGTTCTCGGAAAGTGGAGATCCTCAATGAC 180  
QY 412 GTGATTACCAAGGCGCAGCCCAAGGAAGGCTCCGAGGCGAGGCCAGCTGCCATCTCTATC 471  
Db 181 GTGATTACCAAGGCGCAGCCCAAGGAAGGCTCCGAGGCGAGGCCAGCTGCCATCTCTATC 240  
QY 472 ATCGCCGAGGCTGAGTGTGAGAAATAGCAAGATTCAGCCCCACCTTTTCAGAGCGCAT 531  
Db 241 ATCGCCGAGGCTGAGTGTGAGAAATAGCAAGATTCAGCCCCACCTTTTCAGAGCGCAT 300  
QY 532 TTCAATCGTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAT 591  
Db 301 TTCAATCGTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAT 360  
QY 592 GTGGCCCATGCTACAGAGGGCAAAATGGCCCGTGTGTGGAAAGGGAAGACGCTGCAGC 651  
Db 361 GTGGCCCATGCTACAGAGGGCAAAATGGCCCGTGTGTGGAAAGGGAAGACGCTGCAGC 420  
QY 652 AAAGCCCGGAAGAAACGGAAGAGAGCTCAAAGTCCCTGGCTCATGCGAGGTGGCC 711  
Db 421 AAAGCCCGGAAGAAACGGAAGAGAGCTCAAAGTCCCTGGCTCATGCGAGGTGGCC 480  
QY 712 TTGGCCAAACCCCTCCCGAGGACCCCTGAGCAGAGAGCTGCACCATCCAGTGCAGGAG 771  
Db 481 TTGGCCAAACCCCTCCCGAGGACCCCTGAGCAGAGAGCTGCACCATCCAGTGCAGGAG 540  
QY 772 GATGAGTCTCCAATCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAAGCCTCTG 831  
Db 541 GATGAGTCTCCAATCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAAGCCTCTG 600  
QY 832 AAGGAACAGGCTTTGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCTTACGGCGGCT 891  
Db 601 AAGGAACAGGCTTTGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCTTACGGCGGCT 660  
QY 892 CTGCTCGATCAGAACTCCAAACTGATCAGCCCTTGCATGTCTCAACACACGCTGTGG 951  
Db 661 CTGCTCGATCAGAACTCCAAACTGATCAGCCCTTGCATGTCTCAACACACGCTGTGG 720  
QY 952 AAATGCAACACCCCGAGGCGCCCTTGCCCTGCCCAACGACCCCTTCCCTAT 1011  
Db 721 AAATGCAACACCCCGAGGCGCCCTTGCCCTGCCCAACGACCCCTTCCCTAT 780  
QY 1012 AGCAGATGCTCATCTCCCTTCCATTCACCTCTCCAGCCCTGGAACCTCACCTCTG 1071  
Db 781 AGCAGATGCTCATCTCCCTTCCATTCACCTCTCCAGCCCTGGAACCTCACCTCTG 840  
QY 1072 GAGTCTTCTGGGCAAACTGSCCTGTGTAGACAGCAGAAACCTTGGCTCAGCCACAC 1131  
Db 841 GAGTCTTCTGGGCAAACTGSCCTGTGTAGACAGCAGAAACCTTGGCTCAGCCACAC 900
```


QY 1132 CTGACCAACTGGCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTGGAGGCC 1191
DB 901 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTGGAGGCC 960
QY 1192 AGCTGCCCTGTCTCGTGGTGGCCCATGAGAAGTTTCTGTGGAGGAATACTAGTGCATGCT 1251
DB 961 AGCTGCCCTGTCTCGTGGTGGCCCATGAGAAGTTTCTGTGGAGGAATACTAGTGCATGCT 1020
QY 1252 CTGCAAGGAGCGTGTAGTCTCAAGCCAGCCCAAGCCTGACAGCCTGGCCCAAGACTGG 1311
DB 1021 CTGCAAGGAGCGTGTAGTCTCAAGCCAGCCCAAGCCTGACAGCCTGGCCCAAGACTGG 1080
QY 1312 GCAGACCGGGCTCCAGATCCCGGAGCCGAGCCCAAACTGAGGACAAAGAGGTGTC 1371
DB 1081 GCAGACCGGGCTCCAGATCCCGGAGCCGAGCCCAAACTGAGGACAAAGAGGTGTC 1140
QY 1372 CTGCTCACTGAGAACTCAAGCCAGTGGATTATGATTACCGAGAAAGTCCACTGGGCC 1431
DB 1141 CTGCTCACTGAGAACTCAAGCCAGTGGATTATGATTACCGAGAAAGTCCACTGGGCC 1200
QY 1432 ACGCACAGCTCCGCTCGGAGAGGCTCTTTCGGAGAGGTGCAAGGATGGAGACAAG 1491
DB 1201 ACGCACAGCTCCGCTCGGAGAGGCTCTTTCGGAGAGGTGCAAGGATGGAGACAAG 1260
QY 1492 CAGACTGGCTTCCAGTGGCTGTCAAAGAGTGGCCCTGGAAGTATTTTCGGGAGAGGAG 1551
DB 1261 CAGACTGGCTTCCAGTGGCTGTCAAAGAGTGGCCCTGGAAGTATTTTCGGGAGAGGAG 1320
QY 1552 CTGATGGCATGTGACAGATTGACCTCACCAAGATTGTCCTTTGTATGGAGCTGTGAGA 1611
DB 1321 CTGATGGCATGTGACAGATTGACCTCACCAAGATTGTCCTTTGTATGGAGCTGTGAGA 1380
QY 1612 GAAGGGCTTGGGTCAACATCTTCATGGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1671
DB 1381 GAAGGGCTTGGGTCAACATCTTCATGGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1440
QY 1672 GTCAAGGAGCAGGCTGTCTCCAGAGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTG 1731
DB 1441 GTCAAGGAGCAGGCTGTCTCCAGAGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTG 1500
QY 1732 GAGGTCTGGAATACCTCCACTCAGAAAGATTCTGANTGGGGAGCTCAAAAGCTGACAAAC 1791
DB 1501 GAGGTCTGGAATACCTCCACTCAGAAAGATTCTGANTGGGGAGCTCAAAAGCTGACAAAC 1560
QY 1792 GTGCTCTGTCCAGCGATGGAGGCCACGACCCCTCTGTGACTTTGGCCATGCTGTGTGT 1851
DB 1561 GTGCTCTGTCCAGCGATGGAGGCCACGACCCCTCTGTGACTTTGGCCATGCTGTGTGT 1620
QY 1852 CTTCAACCTGATGGCTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGGCACAGAG 1911
DB 1621 CTTCAACCTGATGGCTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGGCACAGAG 1680
QY 1912 ACCACATGGCTCCGAGGTGGTGTGGGAGGAGCTGCGACGCCAAGGTGGATGTCTGG 1971
DB 1681 ACCACATGGCTCCGAGGTGGTGTGGGAGGAGCTGCGACGCCAAGGTGGATGTCTGG 1740
QY 1972 AGCAGCTGCTGTATGATGCTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTCTTTC 2031
DB 1741 AGCAGCTGCTGTATGATGCTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTCTTTC 1800
QY 2032 CGAGGGCGCTCTGCTCAAGATTGCGAGAGGCTCCGCTGTGAGGGAGATCCCAACC 2091
DB 1801 CGAGGGCGCTCTGCTCAAGATTGCGAGAGGCTCCGCTGTGAGGGAGATCCCAACC 1860
QY 2092 TCCTTGGCCCTCTCAGACCCAGGCCATCAAGAGGGGCTGAGGAAAGGCCCATCCAC 2151
DB 1861 TCCTTGGCCCTCTCAGACCCAGGCCATCAAGAGGGGCTGAGGAAAGGCCCATCCAC 1920
QY 2152 CGCGTGTCTGAGCGGAGCTGGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGGAGGT 2211
DB 1921 CGCGTGTCTGAGCGGAGCTGGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGGAGGT 1980
QY 2212 CTGAAGAGCCCTTGGAGGGGAGATATATAAGAACCAAGACATCCACGCCCAATCAAGCC 2271

DB 1981 CTGAAGAGCCCTTGGAGGGGAGATATATAAGAACCAAGACATCCACGCCCAATCAAGCC 2040
QY 2272 AATTACCAACAGAGCCCTCATGCGCAGCCGAGAGAGCTTTCCGAAGGGCCCCAGGGCCC 2331
DB 2041 AATTACCAACAGAGCCCTCATGCGCAGCCGAGAGAGCTTTCCGAAGGGCCCCAGGGCCC 2100
QY 2332 CGGCCAGCTGAGGAGACAAACAGGAGAGCCCTTAAGTCCAGCCTCTCTCCACACAGAG 2391
DB 2101 CGGCCAGCTGAGGAGACAAACAGGAGAGCCCTTAAGTCCAGCCTCTCTCCACACAGAG 2160
QY 2392 CCCCCAGAGCCAAACAGTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGG 2451
DB 2161 CCCCCAGAGCCAAACAGTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGG 2220
QY 2452 GAACCTTACCTCTGTCTCTCGAGCCAGCCCTGCGAGAAACCCAGCTCACCAGAG 2511
DB 2221 GAACCTTACCTCTGTCTCTCGAGCCAGCCCTGCGAGAAACCCAGCTCACCAGAG 2280
QY 2512 CGGAAAGCAACCGTCCCGAGCAGGAACTGCGAGAGCTGGAAATAGAAATTTCTCTAAC 2571
DB 2281 CGGAAAGCAACCGTCCCGAGCAGGAACTGCGAGAGCTGGAAATAGAAATTTCTCTAAC 2340
QY 2572 AGCCTGTCCCAGCCATTTTCTCTGAGAGCAGAGCAAAATCTCTGTGTCTCAGCATC 2631
DB 2341 AGCCTGTCCCAGCCATTTTCTCTGAGAGCAGAGCAAAATCTCTGTGTCTCAGCATC 2400
QY 2632 GACAGCTCTCTCTGTGCGATGACAGTGAAGAAACCCATCAAAAGGCTCTCAAAAGCTCG 2691
DB 2401 GACAGCTCTCTCTGTGCGATGACAGTGAAGAAACCCATCAAAAGGCTCTCAAAAGCTCG 2460
QY 2692 CGGGACACCTGTAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGCGAGCTCC 2751
DB 2461 CGGGACACCTGTAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGCGAGCTCC 2520
QY 2752 AGCTGGAACTAGTGTCTGGCCCGGGGGGGCCCAACCGACACCCCAAGCTATTTCAATGGT 2811
DB 2521 AGCTGGAACTAGTGTCTGGCCCGGGGGGGCCCAACCGACACCCCAAGCTATTTCAATGGT 2580
QY 2812 GTGAAGTCCAAATACAGTCTCTTAATGTGGAACACCTGACACATCCGGGAGTTCACCGG 2871
DB 2581 GTGAAGTCCAAATACAGTCTCTTAATGTGGAACACCTGACACATCCGGGAGTTCACCGG 2640
QY 2872 GTCAAAAGTGGAGACATCGCCACTGGCATCAGCAGCAGAGTCCAGCTGCGAGCTTCAGC 2931
DB 2641 GTCAAAAGTGGAGACATCGCCACTGGCATCAGCAGCAGAGTCCAGCTGCGAGCTTCAGC 2700
QY 2932 TTGCTACCAAGACGGGCGAGCTTTCGCTAAGCATGAGAGGTGCCAGACTCGGGGCATC 2991
DB 2701 TTGCTACCAAGACGGGCGAGCTTTCGCTAAGCATGAGAGGTGCCAGACTCGGGGCATC 2760
QY 2992 GACCTGAGTGCACATCGGCCCTGATGGGAGCTTCGCTGGAGCTGGAGGCTCAAGCAT 3051
DB 2761 GACCTGAGTGCACATCGGCCCTGATGGGAGCTTCGCTGGAGCTGGAGGCTCAAGCAT 2820
QY 3052 GGCAGCTGGAGAACAGGCCCTTAACCTGCGCTCCACCGCGGCTCCACACTGCGGAAA 3111
DB 2821 GGCAGCTGGAGAACAGGCCCTTAACCTGCGCTCCACCGCGGCTCCACACTGCGGAAA 2880
QY 3112 GCAGCTTCTGTCTGTGTGCAAGATGTGCTGCTGAAACACAGGCTCAGCGCTTCACAGG 3171
DB 2881 GCAGCTTCTGTCTGTGTGCAAGATGTGCTGCTGAAACACAGGCTCAGCGCTTCACAGG 2940
QY 3172 GGATTTGCGCCCGCCCGGCTCAGTGGGAAACCGAGGCTTCGAGCAGCAGAGTGGGGGC 3231
DB 2941 GGATTTGCGCCCGCCCGGCTCAGTGGGAAACCGAGGCTTCGAGCAGCAGAGTGGGGGC 3000
QY 3232 AAGCAGAAATGCTCCAGGATTTTACACCTGAGGCTTCGCCACCTGCTGAAAAAACAT 3291
DB 3001 AAGCAGAAATGCTCCAGGATTTTACACCTGAGGCTTCGCCACCTGCTGAAAAAACAT 3060
QY 3292 CCGCCACGTGAAGAGACAGAGGAGGTGCGAGAGTTTACTTGGGGAAACAAAAACAGGA 3351

Db 3061 CGCCACGCTGAAGACACAGAAGGAGGATGCGAGGATTACCTGGGAAACAAACACAGGA 3120

Qy 3352 TCTTTTCTGCCCTGCTCCAGTCCAGTTGGCTTGA 3387

Db 3121 TCTTTTCTGCCCTGCTCCAGTCCAGTTGGCTTGA 3156

RESULT 8

AR063065

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Query Match 68.5%; Score 3149.6; DB 6; Length 3156;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 232 ATGGCAGTATGAAATGSCCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAG 291

Db 1 ATGGCAGTATGAAATGSCCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAG 60

Qy 292 GAACTCCCAAGCCAAAGAGAGAGCGCCGCACTCGGGGAAGAAACAGAGCTCCGTCTAC 351

Db 61 GAACTCCCAAGCCAAAGAGAGAGCGCCGCACTCGGGGAAGAAACAGAGCTCCGTCTAC 120

Qy 352 AAGCTTGAGCCGTGAGAGAGCCCTGTCTGCGGAAGTGGAGATCCCTGAATGAC 411

Db 121 AAGCTTGAGCCGTGAGAGAGCCCTGTCTGCGGAAGTGGAGATCCCTGAATGAC 180

Qy 412 GTGATTACCAAGGCGCAGCCAGGAGGCTCCGAGGCGAGGCGAGCTGCCATCTCTATC 471

Db 181 GTGATTACCAAGGCGCAGCCAGGAGGCTCCGAGGCGAGGCGAGCTGCCATCTCTATC 240

Qy 472 ATCGCCAGGCTGAGTGTGAGAAATAGCCAGATTGAGCCCACTTTTTCAGAAACGCAAT 531

Db 241 ATCGCCAGGCTGAGTGTGAGAAATAGCCAGATTGAGCCCACTTTTTCAGAAACGCAAT 300

Qy 532 TTGATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGGCTTGTATCAGATCCCAACAAT 591

Db 301 TTGATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGGCTTGTATCAGATCCCAACAAT 360

Qy 592 GTGGCCCATCTACAGAGGCAAAATGGCCCGTGTGTGTGGAAAGGAAAGCGTCGACG 651

Db 361 GTGGCCCATCTACAGAGGCAAAATGGCCCGTGTGTGTGGAAAGGAAAGCGTCGACG 420

Qy 652 AAAGCCCGGAAGAAACGGAAGAGAGAGTCTAAAGTCCCTGGCTCATGAGAGTGGCC 711

Db 421 AAAGCCCGGAAGAAACGGAAGAGAGAGTCTAAAGTCCCTGGCTCATGAGAGTGGCC 480

Qy 712 TTGGCCAAACCCCTCCCAAGGACCCCTGAGCAGAGAGTGCACCATCCCAAGTGCAGGAG 771

Db 481 TTGGCCAAACCCCTCCCAAGGACCCCTGAGCAGAGAGTGCACCATCCCAAGTGCAGGAG 540

Qy 772 GATGAGTCTCCACTCGCGGCCCATATGTTAGAAACACCCCGAGTTCACCAAGCTCTG 831

Db 541 GATGAGTCTCCACTCGCGGCCCATATGTTAGAAACACCCCGAGTTCACCAAGCTCTG 600

Qy 832 AAGGAACAGGCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCTTACGCGCGCT 891

Db 601 AAGGAAACGAGCCCTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGCGCTACGCGCGCT 660

Qy 892 CTGCTCGATCAGAACTCCACAAACTGATCAGGCCCCCTTGCAATGTCTGTGAACCAACGTTGG 951

Db 661 CTGCTCGATCAGAACTCCACAAACTGATCAGGCCCCCTTGCAATGTCTGTGAACCAACGTTGG 720

Qy 952 AAACTGCAACACCCCGAGGAGCGGAGCCCTGCGCTGCGCCAGGCAACCCCTTCCCTAT 1011

Db 721 AAACTGCAACACCCCGAGGAGCGGAGCCCTGCGCTGCGCCAGGCAACCCCTTCCCTAT 780

Qy 1012 AGCAGACTGCTCATCCCTTCCCATTCACCTCTCCAGCCCTTGGAAACCTCACCCTCTG 1071

Db 781 AGCAGACTGCTCATCCCTTCCCATTCACCTCTCCAGCCCTTGGAAACCTCACCCTCTG 840

Qy 1072 GAGTCTCTCTGGGCAAACTGGGCTGTGTAGACAGCCAGCAAAACCTTGGCTGTGACCCAC 1131

Db 841 GAGTCTCTCTGGGCAAACTGGGCTGTGTAGACAGCCAGCAAAACCTTGGCTGTGACCCAC 900

Qy 1132 CTGAGCAAACTGGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTTGGAGCCC 1191

Db 901 CTGAGCAAACTGGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTTGGAGCCC 960

Qy 1192 AGCTGCTGTCTGCTGCTGCCATGAGAGTCTTCTGTGAGGGAATACCTAGTGCATGCT 1251

Db 961 AGCTGCTGTCTGCTGCTGCCATGAGAGTCTTCTGTGAGGGAATACCTAGTGCATGCT 1020

Qy 1252 CTGCAAGCGAGCTGAGCTCAAGCCAGGCCACAGCCTGACAGCCTGGGCAAGACCTGG 1311

Db 1021 CTGCAAGCGAGCTGAGCTCAAGCCAGGCCACAGCCTGACAGCCTGGGCAAGACCTGG 1080

Qy 1312 GCAGCAGGGGCTCCAGATCCCGGAGCCAGCCCCCAAACTGAGGCAACAGGGGTGTC 1371

Db 1081 GCAGCAGGGGCTCCAGATCCCGGAGCCAGCCCCCAAACTGAGGCAACAGGGGTGTC 1140

Qy 1372 CTGCTCACTGAGAACTCAAGCCAGTGGATTATCAGTACCGAGAGAGTCCACTGGGSC 1431

Db 1141 CTGCTCACTGAGAACTCAAGCCAGTGGATTATCAGTACCGAGAGAGTCCACTGGGSC 1200

Qy 1432 ACGCACAGCTCCGCTGGGCGAGAGGCTCTTCGAGAGGTGCAAGGATGGAGACAAG 1491

Db 1201 ACGCACAGCTCCGCTGGGCGAGAGGCTCTTCGAGAGGTGCAAGGATGGAGACAAG 1260

Qy 1492 CAGACTGCTTCCAGTGGGCTGCAAAAGGTCGCGCTGGAAGTATTTTCGGGCGAGAGAG 1551

Db 1261 CAGACTGCTTCCAGTGGGCTGCAAAAGGTCGCGCTGGAAGTATTTTCGGGCGAGAGAG 1320

Qy 1552 CTGATGGCATGTGAGAGATTGACCTCAACCCAGAAATTTGCTTGTATGAGAGTGTGAGA 1611

Db 1321 CTGATGGCATGTGAGAGATTGACCTCAACCCAGAAATTTGCTTGTATGAGAGTGTGAGA 1380

Qy 1612 GAAGGCGCTTGGGTCAACATCTTCATGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1671

Db 1381 GAAGGCGCTTGGGTCAACATCTTCATGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1440

Qy 1672 GTCAAGGAGCAGGCTGTCTCCAGAGACCGGGCCGTGTACTACTCTGGGCGAGGCCCTG 1731

Db 1441 GTCAAGGAGCAGGCTGTCTCCAGAGACCGGGCCGTGTACTACTCTGGGCGAGGCCCTG 1500

Qy 1732 GAGGCTCTGGAATACCTCCACTCAAGAAAGATTCTGCTGAGGAGCGTCAAAAGCTCAAC 1791

Db 1501 GAGGCTCTGGAATACCTCCACTCAAGAAAGATTCTGCTGAGGAGCGTCAAAAGCTCAAC 1560

Qy 1792 GTGCTCTGTCCAGCGATGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGT 1851

Db 1561 GTGCTCTGTCCAGCGATGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1620

Qy 1852 CTTCAACCTGATGGCTGGGAAAGTCTTGTCTCAGAGGAGCTACATCCCTGGCAGAG 1911

Db 1621 CTTCAACCTGATGGCTGGGAAAGTCTTGTCTCAGAGGAGCTACATCCCTGGCAGAG 1680

Qy 1912 ACCCAGATGGCTCCGAGAGTGTGTGGCAGGAGCTGCGACGCCAAGGTGATGCTGTG 1971

Db 1681 ACCCAGATGGCTCCGAGAGTGTGTGGCAGGAGCTGCGACGCCAAGGTGATGCTGTG 1740

QY 592 GTGGCCCATCTTACAGAGGCAAAATGGCCCGTGTGTGTGGAGAGGAAAGCGTCGCAGC 651
DB |||||
QY 361 GTGGCCCATCTTACAGAGGCAAAATGGCCCGTGTGTGTGGAGAGGAAAGCGTCGCAGC 420
DB |||||
QY 652 AAAGCCCGAAGAAAACGGAAGAGAGCTCAAAAGTCCCTGGCTCATGCAGGAGTGCC 711
DB |||||
QY 421 AAAGCCCGAAGAAAACGGAAGAGAGCTCAAAAGTCCCTGGCTCATGCAGGAGTGCC 480
DB |||||
QY 712 TTGGCCAAACCCCTCCAGAGCCCTGAGCAGAGAGCTGCACCATCCAGTGCAGGAG 771
DB |||||
QY 481 TTGGCCAAACCCCTCCAGAGCCCTGAGCAGGAGAGCTGCACCATCCAGTGCAGGAG 540
DB |||||
QY 772 GATGAGTCTCCACTCGGGCCCAATATGTTAGAAAACCCCGCAGTTCACCAAGCCTCTG 831
DB |||||
QY 541 GATGAGTCTCCACTCGGGCCCAATATGTTAGAAAACCCCGCAGTTCACCAAGCCTCTG 600
DB |||||
QY 832 AAGAAACCAAGGCTTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGGCTACGGCCGGCT 891
DB |||||
QY 601 AAGAAACCAAGGCTTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGGCTACGGCCGGCT 660
DB |||||
QY 892 CTGCTCGATCAGAACTCCAAAACCTGATCAGCCCTTGCATGCTGAAACCAACGCTGGG 951
DB |||||
QY 661 CTGCTCGATCAGAACTCCAAAACCTGATCAGCCCTTGCATGCTGAAACCAACGCTGGG 720
DB |||||
QY 952 AAACCTGCACACCCCGAGAGCGAGGCCCTGCCCCCTGCCACACGCCCTTCCCTAT 1011
DB |||||
QY 721 AAACCTGCACACCCCGAGAGCGAGGCCCTGCCCCCTGCCACACGCCCTTCCCTAT 780
DB |||||
QY 1012 AGCAGATGCTCATCCCTTCCCATTTCCACCCCTCTCCAGCCCTGGAACCTCACCCCTCTG 1071
DB |||||
QY 781 AGCAGATGCTCATCCCTTCCCATTTCCACCCCTCTCCAGCCCTGGAACCTCACCCCTCTG 840
DB |||||
QY 1072 GAGTCTTCTGGCAAACTGGCTGTAGACAGTCCAAAGCCCTGCTGGGCCACACCTGGAGCCC 1131
DB |||||
QY 841 GAGTCTTCTGGCAAACTGGCTGTAGACAGTCCAAAGCCCTGCTGGGCCACACCTGGAGCCC 900
DB |||||
QY 1132 CTGAGCAAACTGGCTGTAGACAGTCCAAAGCCCTGCTGGGCCACACCTGGAGCCC 1191
DB |||||
QY 901 CTGAGCAAACTGGCTGTAGACAGTCCAAAGCCCTGCTGGGCCACACCTGGAGCCC 960
DB |||||
QY 1192 AGCTGCTGTCTCGTGGTCCCATGAGAAATTTCTGTGGAGGAATACCTAGTGAATGCT 1251
DB |||||
QY 961 AGCTGCTGTCTCGTGGTCCCATGAGAAATTTCTGTGGAGGAATACCTAGTGAATGCT 1020
DB |||||
QY 1252 CTGCAAGGACGGTGAATCTCAAGCCAGGCCCAAGCTTACAGCTTGGCCCAAGACCTGG 1311
DB |||||
QY 1021 CTGCAAGGACGGTGAATCTCAAGCCAGGCCCAAGCTTACAGCTTGGCCCAAGACCTGG 1080
DB |||||
QY 1312 GCAGACGGGGCTCCAGATCCCGGAGCCAGCCCCCAAACTGAGGACAAACGAGGGTGTCT 1371
DB |||||
QY 1081 GCAGACGGGGCTCCAGATCCCGGAGCCAGCCCCCAAACTGAGGACAAACGAGGGTGTCT 1140
DB |||||
QY 1372 CTGCTCACTGAGAACTCAAGCCAGTGGATATGAGTACCGAGAGAAAGTCCACTGGGCC 1431
DB |||||
QY 1141 CTGCTCACTGAGAACTCAAGCCAGTGGATATGAGTACCGAGAGAAAGTCCACTGGGCC 1200
DB |||||
QY 1432 AGCCACAGCTCGGCTGGGACAGGCTCTTGGAGAGTGCACAGGATGAGGACAG 1491
DB |||||
QY 1201 AGCCACAGCTCGGCTGGGACAGGCTCTTGGAGAGGTGCACAGGATGAGGACAG 1260
DB |||||
QY 1492 CAGACTGGCTTCAGTGGCTGTCAAAAAGTGGCGCTGGAAATATTTTCGGGCAGAGGAG 1551
DB |||||
QY 1261 CAGACTGGCTTCAGTGGCTGTCAAAAAGTGGCGCTGGAAATATTTTCGGGCAGAGGAG 1320
DB |||||
QY 1552 CTGATGGCATGTGCAGGATGACCTCAACCAAGATGCTCCCTTTGATGAGCTGTGAGA 1611
DB |||||
QY 1321 CTGATGGCATGTGCAGGATGACCTCAACCAAGATGCTCCCTTTGATGAGCTGTGAGA 1380
DB |||||
QY 1612 GAAGGCCCTTGGGTCAACATCTTCATGAGAGCTCTGGAGGTGGCTCCCTGGGCCAGCTG 1671
DB |||||
QY 1381 GAAGGCCCTTGGGTCAACATCTTCATGAGAGCTCTGGAGGTGGCTCCCTGGGCCAGCTG 1440
DB |||||
QY 1672 GTCAAGGACAGGGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGGCCCTG 1731
DB |||||

DB |||||
QY 1441 GTCAAGGACAGGGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGCCCTG 1500
DB |||||
QY 1732 GAGGGTCTGGAAATACCTCCACTCAAGAGGATTTCTGCATGGGGACGTCAAAGCTGCACAC 1791
DB |||||
QY 1501 GAGGGTCTGGAAATACCTCCACTCAAGAGGATTTCTGCATGGGGACGTCAAAGCTGCACAC 1560
DB |||||
QY 1792 GTGCTCCTGTCCAGCGATGGGAGCCAGCGACCCCTCTGTGACTTTGGCCATGCTGTGTGT 1851
DB |||||
QY 1561 GTGCTCCTGTCCAGCGATGGGAGCCAGCGACCCCTCTGTGACTTTGGCCATGCTGTGTGT 1620
DB |||||
QY 1852 CTTCAACCTGATGCGCTGGGAAAGTCTTGTCTCAAGGGACTACATCCCTGGGCACAGAG 1911
DB |||||
QY 1621 CTTCAACCTGATGCGCTGGGAAAGTCTTGTCTCAAGGGACTACATCCCTGGGCACAGAG 1680
DB |||||
QY 1912 ACCCACATGSGCTCGGAGGTGCTGCGGAGGAGCTGCGACGCCCAAGGTGATGTCTGG 1971
DB |||||
QY 1681 ACCCACATGSGCTCGGAGGTGCTGCGGAGGAGCTGCGACGCCCAAGGTGATGTCTGG 1740
DB |||||
QY 1972 AGCAGCTGCTATGATGCTGCATGCTCAACCGCTGCCACCCCTGAGCTCAGTTCTTC 2031
DB |||||
QY 1741 AGCAGCTGCTATGATGCTGCATGCTCAACCGCTGCCACCCCTGAGCTCAGTTCTTC 1800
DB |||||
QY 2032 CGAGGGCGCTCTGCTCTCAAGATTGCGAGCGCTCCGCTGTGAGGGAGATCCACCC 2091
DB |||||
QY 1801 CGAGGGCGCTCTGCTCTCAAGATTGCGAGCGCTCCGCTGTGAGGGAGATCCACCC 1860
DB |||||
QY 2092 TCCTGGCGCCCTCTCAAGCGCCAGGCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCAC 2151
DB |||||
QY 1861 TCCTGGCGCCCTCTCAAGCGCCAGGCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCAC 1920
DB |||||
QY 2152 CGCTGTCTGAGCGGAGCTGGGAGGAGTGAACCGGGCACTACAGCAAGTGGAGGT 2211
DB |||||
QY 1921 CGCTGTCTGAGCGGAGCTGGGAGGAGTGAACCGGGCACTACAGCAAGTGGAGGT 1980
DB |||||
QY 2212 CTGAAGAGCCCTTGGAGGGGAGATATAAGAAACCAAGCATCCACCGCCCAAAATCAAGCC 2271
DB |||||
QY 1981 CTGAAGAGCCCTTGGAGGGGAGATATAAGAAACCAAGCATCCACCGCCCAAAATCAAGCC 2040
DB |||||
QY 2272 AATTACACACAGACCTTCATGCCCGCGAGAGAGCTTCGCCAAGGGCCCGCAGGGCCC 2331
DB |||||
QY 2041 AATTACACACAGACCTTCATGCCCGCGAGAGAGCTTCGCCAAGGGCCCGCAGGGCCC 2100
DB |||||
QY 2332 CGGCCAGCTGAGGAGACAAACAGGACAGCCCTCTAAGCTCCAGCCCTCTCTCCACACAGAG 2391
DB |||||
QY 2101 CGGCCAGCTGAGGAGACAAACAGGACAGCCCTCTAAGCTCCAGCCCTCTCTCCACACAGAG 2160
DB |||||
QY 2392 CCCCAGAGCCAAACAAAGTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGG 2451
DB |||||
QY 2161 CCCCAGAGCCAAACAAAGTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGG 2220
DB |||||
QY 2452 GAAACCTTTACCTCTGTCTCTCCCTGGAGCCAGCCCTGCGAGAAACCCAGAGCTCACAGAG 2511
DB |||||
QY 2221 GAAACCTTTACCTCTGTCTCTCCCTGGAGCCAGCCCTGCGAGAAACCCAGAGCTCACAGAG 2280
DB |||||
QY 2512 CGGAAAGCAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAATTTCTCTCAAC 2571
DB |||||
QY 2281 CGGAAAGCAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAATTTCTCTCAAC 2340
DB |||||
QY 2572 AGCTGTCTCCAGCCATTTTCTGTGAGGAGCAGAGCAAAATTTCTCTGTGCTCTCAGATC 2631
DB |||||
QY 2341 AGCTGTCTCCAGCCATTTTCTGTGAGGAGCAGAGCAAAATTTCTCTGTGCTCTCAGATC 2400
DB |||||
QY 2632 GACAGCTCTCCCTGTCCGATCAGAGTGCAGAGAACCCATCAAGAGCTCTCAAAGCTCG 2691
DB |||||
QY 2401 GACAGCTCTCCCTGTCCGATCAGAGTGCAGAGAACCCATCAAGAGCTCTCAAAGCTCG 2460
DB |||||
QY 2692 CGGAGACCCCTGAGCTCAGGGCTACATCTCTGGAGCAGCAGGCCCGAGGCTCGAAGCTCC 2751
DB |||||
QY 2461 CGGAGACCCCTGAGCTCAGGGCTACATCTCTGGAGCAGCAGGCCCGAGGCTCGAAGCTCC 2520
DB |||||
QY 2752 AGCTGGAACATGCTGTGCTCGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGCT 2811
DB |||||

```
Db 2521 AGCTGGAACATGGTCTGCTGGCGGGGGGGCCACCGACACCCAGACTATTTCAATGGT 2580
Qy 2812 GTGAAAGTCCAAATAACAGTCTCTTAATAGTGAACAACCTGCACATCCGGAGTTCACACGG 2871
Db 2581 GTGAAAGTCCAAATAACAGTCTCTTAATAGTGAACAACCTGCACATCCGGAGTTCACACGG 2640
Qy 2872 GTCAAGTGGGAGACATCCCACTGGGATCAGCAGCAGATCCCACTGCAAGCTTCAGC 2931
Db 2641 GTCAAGTGGGAGACATCCCACTGGGATCAGCAGCAGATCCCACTGCAAGCTTCAGC 2700
Qy 2932 TTGGTCAACAAAGACGGGAGCCTGTTCGTACGACATGGAGGTGCCAGACTCGGGCATC 2991
Db 2701 TTGGTCAACAAAGACGGGAGCCTGTTCGTACGACATGGAGGTGCCAGACTCGGGCATC 2760
Qy 2992 GACCTGCAGTGCACATGCGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGGTCAAGCAT 3051
Db 2761 GACCTGCAGTGCACATGCGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGGTCAAGCAT 2820
Qy 3052 GGCCAGCTGGAGACAGCCCTTAACCTGCTCCACCGCGGCTCCACATGCGCGGAA 3111
Db 2821 GGCCAGCTGGAGACAGCCCTTAACCTGCTCCACCGCGGCTCCACATGCGCGGAA 2880
Qy 3112 GCAGCTTCCTGCTCGGTGACGATGCTGCTGAAACACAGGCTCAGCCGTTCCAGG 3171
Db 2881 GCAGCTTCCTGCTCGGTGACGATGCTGCTGAAACACAGGCTCAGCCGTTCCAGG 2940
Qy 3172 GGATTGCCAGCCCCCGCTCACAGTGGGAACAGGGCTCGCAGCAGCAAGTGGGGC 3231
Db 2941 GGATTGCCAGCCCCCGCTCACAGTGGGAACAGGGCTCGCAGCAGCAAGTGGGGC 3000
Qy 3232 AAGCAGAAATGCTCCAGGATTCACCTCAGCCCTGCCCCCAGCCCTGCTGAAAAACAT 3291
Db 3001 AAGCAGAAATGCTCCAGGATTCACCTCAGCCCTGCCCCCAGCCCTGCTGAAAAACAT 3060
Qy 3292 CCGCCAGCTGAAGACAGAGAGGAGTGGCAGGATTAACCTGGGGAACAAACAGGA 3351
Db 3061 CCGCCAGCTGAAGACAGAGAGGAGTGGCAGGATTAACCTGGGGAACAAACAGGA 3120
Qy 3352 TCTTTTCTGCCCCCTGCTCCAGTCCAGTTCAGTTGGCTGA 3387
Db 3121 TCTTTTCTGCCCCCTGCTCCAGTTCAGTTGGCTGA 3156

RESULT 10
LOCUS BD062401 3156 bp DNA linear PAT 27-AUG-2002
DEFINITION NIK proteins, nucleic acids and methods.
ACCESSION BD062401
VERSION BD062401.1 GI:22608004
KEYWORDS JP 2001510348-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Rothe.M. and Wu.L.
NIK proteins, nucleic acids and methods
Patent: JP 2001510348-A 1 31-JUL-2001;
TULARIK INC
PN JP 2001510348-A/1
PD 31-JUL-2001
PF 02-JUL-1998 JP 1999507409
PR 03-JUL-1997 US 08/887518
PI MIKE ROTHE, LIN WU
PC
C07K7/06, C07K7/08, C07K14/47, C12N1/00, C12N5/10, C12N5/12, C12N15/ PC
63,
PC C12P21/02, G01N33/53
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
1. .3156
```

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 68.5%; Score 3149.6; DB 6; Length 3156;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 232 ATGCAGTGTATGGAATAGGCTGCGCAGGTGCGCCCTGCTCAGCAGTGGGCGAGCAGAG 291
Db 1 ATGCAGTGTATGGAATAGGCTGCGCAGGTGCGCCCTGCTCAGCAGTGGGCGAGCAGAG 60
Qy 292 GAATCTCCCAAGCCAAAGGAGAGACGCCGCCACTGCTGGGGAAGAAACAGAGTCCGTCTAC 351
Db 61 GAATCTCCCAAGCCAAAGGAGAGACGCCGCCACTGCTGGGGAAGAAACAGAGTCCGTCTAC 120
Qy 352 AAGCTTTCAGGCGGTGGAGAGAGAGCCCTGTGTTCTGCGGAAGTGGAGATCCTGAATGAC 411
Db 121 AAGCTTTCAGGCGGTGGAGAGAGAGCCCTGTGTTCTGCGGAAGTGGAGATCCTGAATGAC 180
Qy 412 GTGATTACCAAGGACACAGCAAGGAAGGCTCCAGGCGAGGCGCAGCTGCGCATCTCTATC 471
Db 181 GTGATTACCAAGGACACAGCAAGGAAGGCTCCAGGCGAGGCGCAGCTGCGCATCTCTATC 240
Qy 472 ATCCGCCAGGCTGAGTGTGAGAATAGCAAGAGTTTCAGCCCCACCTTTTCAGAAAGCAT 531
Db 241 ATCCGCCAGGCTGAGTGTGAGAATAGCAAGAGTTTCAGCCCCACCTTTTCAGAAAGCAT 300
Qy 532 TTCTATCCTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAT 591
Db 301 TTCTATCCTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAT 360
Qy 592 GTGCCCATGCTACAGAGGGCAAAATGCGCCGTGTGTTGGAAGGAAAGCGTCGCAGC 651
Db 361 GTGCCCATGCTACAGAGGGCAAAATGCGCCGTGTGTTGGAAGGAAAGCGTCGCAGC 420
Qy 652 AAAGCCCGAAGAAACCGAAGAGAGTCAAGTCCCTGGCTCATGCAAGGAGTGGCC 711
Db 421 AAAGCCCGAAGAAACCGAAGAGAGTCAAGTCCCTGGCTCATGCAAGGAGTGGCC 480
Qy 712 TTGCCCAAAACCCCTCCACAGGACCCCTGAGCAGAGAGTGCACATCCAGTGGAGAG 771
Db 481 TTGCCCAAAACCCCTCCACAGGACCCCTGAGCAGAGAGTGCACATCCAGTGGAGAG 540
Qy 772 GATGAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCCTCTG 831
Db 541 GATGAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCCTCTG 600
Qy 832 AAGGAACCCAGGCTTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGCTTACGGCCGGCT 891
Db 601 AAGGAACCCAGGCTTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGCTTACGGCCGGCT 660
Qy 892 CTGCTCGATCAGAACTCCAAACTGATCAGCCCTTGCATGTCGAAACACAGTGTGG 951
Db 661 CTGCTCGATCAGAACTCCAAACTGATCAGCCCTTGCATGTCGTAACACAGTGTGG 720
Qy 952 AAACCTGACACACCCCGCAGGACGGAGCCCTGCGCCCTGCGCCACGACCCCTTCCCTTAT 1011
Db 721 AAACCTGACACACCCCGCAGGACGGAGCCCTGCGCCCTGCGCCACGACCCCTTCCCTTAT 780
Qy 1012 AGCAGACTGCTCATCTCCCTTCCATTCACCCCTCTCCAGCCCTTGGAAACCTTCACCCCTCTG 1071
Db 781 AGCAGACTGCTCATCTCCCTTCCATTCACCCCTCTCCAGCCCTTGGAAACCTTCACCCCTCTG 840
Qy 1072 GAGTCTCTCTGGGCAAACTGGGCTGTGTAGACAGCAGAAACCCCTTGCCTGACCCACAC 1131
Db 841 GAGTCTCTCTGGGCAAACTGGGCTGTGTAGACAGCAGAAACCCCTTGCCTGACCCACAC 900
Qy 1132 CTGAGCAAACTGGGCTGTGTAGACAGTCCAAAGCCCTTGCCTGCGCCACACCTGGAGGCC 1191
Db 901 CTGAGCAAACTGGGCTGTGTAGACAGTCCAAAGCCCTTGCCTGCGCCACACCTGGAGGCC 960
```

QY 1192 AGCTGCTGCTCGTGTGCGCCATCAGAAAGTTTTCTGTGGAGGATACCTAGTGCATGCT 1251
DB 961 AGCTGCTGCTCGTGTGCGCCATCAGAAAGTTTTCTGTGGAGGATACCTAGTGCATGCT 1020
QY 1252 CTGCAAGGAGCGGTGAGCTCAAGCCAGGCCCAAGCCTGACAGCCTGGCCAAAGCCTGG 1311
DB 1021 CTGCAAGGAGCGGTGAGCTCAAGCCAGGCCCAAGCCTGACAGCCTGGCCAAAGCCTGG 1080
QY 1312 GCAGCAGCGGGCTCAGATCCCGGGAGGCCAGCCCCCAAACTGAGGACAAAGAGGGTGTG 1371
DB 1081 GCAGCAGCGGGCTCAGATCCCGGGAGGCCAGCCCCCAAACTGAGGACAAAGAGGGTGTG 1140
QY 1372 CTGCTCACTGAGAAACTCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCC 1431
DB 1141 CTGCTCACTGAGAAACTCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCC 1200
QY 1432 ACGCACCAAGCTCCGCTCGGCGAGAGGCTCTCTCGAGAGGTGCA CAGATGAGAGCAAG 1491
DB 1201 ACGCACCAAGCTCCGCTCGGCGAGAGGCTCTCTCGAGAGGTGCA CAGATGAGAGCAAG 1260
QY 1492 CAGACTGGCTTCAGTGGCGCTGTCAAAAGGTGCGCTGGAAGTATTTGGGCGAGAGG 1551
DB 1261 CAGACTGGCTTCAGTGGCGCTGTCAAAAGGTGCGCTGGAAGTATTTGGGCGAGAGG 1320
QY 1552 CTGATGGCATGTGCAGGATTGACCTCACCCAGAAATGTCCCTTTGTATGGAGCTGTGAGA 1611
DB 1321 CTGATGGCATGTGCAGGATTGACCTCACCCAGAAATGTCCCTTTGTATGGAGCTGTGAGA 1380
QY 1612 GAAGGCGCTTGGGTCAACATCTTCATGAGCTGTGGAAGTGGTTCCTTGGGCCAGCTG 1671
DB 1381 GAAGGCGCTTGGGTCAACATCTTCATGAGCTGTGGAAGTGGTTCCTTGGGCCAGCTG 1440
QY 1672 GTCCAGGACAGGGCTGTCTCCAGAGGACCGGGCCCTGTACTCTGGGCCAGGCCCTG 1731
DB 1441 GTCCAGGACAGGGCTGTCTCCAGAGGACCGGGCCCTGTACTCTGGGCCAGGCCCTG 1500
QY 1732 GAGGGTCTGGAATACCTCCACCTCACGAAGGATTTCTGCATGGGAGCTCAAAAGCTGACAAC 1791
DB 1501 GAGGGTCTGGAATACCTCCACCTCACGAAGGATTTCTGCATGGGAGCTCAAAAGCTGACAAC 1560
QY 1792 GTGCTCTCTCCAGCGATGGGAGCCACGAGCCCTCTGTGACTTTGGGCCATGCTGTGTG 1851
DB 1561 GTGCTCTCTCCAGCGATGGGAGCCACGAGCCCTCTGTGACTTTGGGCCATGCTGTGTG 1620
QY 1852 CTTCAACCTGATGGCTGGGAAGTCTTGTCTCAGAGGGACTACATCCCTGGCAGAGAG 1911
DB 1621 CTTCAACCTGATGGCTGGGAAGTCTTGTCTCAGAGGGACTACATCCCTGGCAGAGAG 1680
QY 1912 ACCCATGAGCTCCGAGGTGGTGTGGGAGAGCTGCGAGAGCTGCGAGGAGTGTCTGG 1971
DB 1681 ACCCATGAGCTCCGAGGTGGTGTGGGAGAGCTGCGAGAGCTGCGAGGAGTGTCTGG 1740
QY 1972 AGCAGTGTGTATGATGTGCATGTCTCAACGGCTGCGACCCCTGGACTCAGTTCTTTC 2031
DB 1741 AGCAGTGTGTATGATGTGCATGTCTCAACGGCTGCGACCCCTGGACTCAGTTCTTTC 1800
QY 2032 CGAGGCGCTGTGCTCAGATTGCGAGGAGCTCCGCTGTGAGGAGATCCCAACC 2091
DB 1801 CGAGGCGCTGTGCTCAGATTGCGAGGAGCTCCGCTGTGAGGAGATCCCAACC 1860
QY 2092 TCTGCGCCCTCTCACAGCCAGGCCATCCAAAGAGGGCTGAGAAAGAGGCCATCCAC 2151
DB 1861 TCTGCGCCCTCTCACAGCCAGGCCATCCAAAGAGGGCTGAGAAAGAGGCCATCCAC 1920
QY 2152 CGCGTGTGTGACGGAGCTGGAGGGAAAGGTGAAACGGGCACTACAGCAAGTGGAGGT 2211
DB 1921 CGCGTGTGTGACGGAGCTGGAGGGAAAGGTGAAACGGGCACTACAGCAAGTGGAGGT 1980
QY 2212 CTGAAGGCCCTTGGAGGGAGATATAAGAACCAAGACATCCACCGCCCAATCAAGCC 2271
DB 1981 CTGAAGGCCCTTGGAGGGAGATATAAGAACCAAGACATCCACCGCCCAATCAAGCC 2040
QY 2272 AATTACCACAGACCCCTCCATGTCACGCCAGAGAGCTTTTCGCAAGGGCCCCAGGGCCC 2331

DB 2041 AATTACCACAGACCCCTCCATGTCACGCCAGAGAGCTTTTCGCAAGGGCCCCAGGGCCC 2100
QY 2332 CGGCCAGCTGAGGAGACAAAGCAGAGCCCTAAGCTCAGCCCTCTCTCTCCCAACAGAG 2391
DB 2101 CGGCCAGCTGAGGAGACAAAGCAGAGCCCTAAGCTCAGCCCTCTCTCTCCCAACAGAG 2160
QY 2392 CCCCAGAGCCAAACAAAGTCTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTG 2451
DB 2161 CCCCAGAGCCAAACAAAGTCTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTG 2220
QY 2452 GAAACCTTACTCTGTCTCTCTGAGCAGCCCTGCGAGAAACCCAGCTCAGCAGAG 2511
DB 2221 GAAACCTTACTCTGTCTCTCTGAGCAGCCCTGCGAGAAACCCAGCTCAGCAGAG 2280
QY 2512 CGGAAAGCAACCGTCCCGGAGCAGAACTGCAGCAGCTGGAAATAGAAATATTCTCTCAAC 2571
DB 2281 CGGAAAGCAACCGTCCCGGAGCAGAACTGCAGCAGCTGGAAATAGAAATATTCTCTCAAC 2340
QY 2572 AGCCTGTCCCAGCCATTTCTCTGGAGGAGCAGAGCAAAATTTCTCTGTCCTCAGCATC 2631
DB 2341 AGCCTGTCCCAGCCATTTCTCTGGAGGAGCAGAGCAAAATTTCTCTGTCCTCAGCATC 2400
QY 2632 GACAGCCTCTCTCTGTGCGATGACAGTGAAGAAACCCATCAAAAGGCTCTCTCAAGCTCG 2691
DB 2401 GACAGCCTCTCTCTGTGCGATGACAGTGAAGAAACCCATCAAAAGGCTCTCTCAAGCTCG 2460
QY 2692 CGGCAACCCCTGAGCTCAGGGGTACACTCTCTGGAGCAGCAGGCGGAGCTCGAAGCTCC 2751
DB 2461 CGGCAACCCCTGAGCTCAGGGGTACACTCTCTGGAGCAGCAGGCGGAGGCTCGAAGCTCC 2520
QY 2752 AGCTGGAACTGTGTGTCGGCCCGGGCGGCCACACGCCAAGCTATTTCAATGGT 2811
DB 2521 AGCTGGAACTGTGTGTCGGCCCGGGCGGCCACACGCCAAGCTATTTCAATGGT 2580
QY 2812 GTGAAAGTCCAAATACAGTCTCTTAATGTGTAACAACCTGACATCCGGGAGTTCCACCG 2871
DB 2581 GTGAAAGTCCAAATACAGTCTCTTAATGTGTAACAACCTGACATCCGGGAGTTCCACCG 2640
QY 2872 GTCAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCGAGCCTTCAAGC 2931
DB 2641 GTCAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCGAGCCTTCAAGC 2700
QY 2932 TTGGTCAACAAAGACGGGACGCTGTTGCTACGACATGGAGGTGCCAGATCTCGGGCATC 2991
DB 2701 TTGGTCAACAAAGACGGGACGCTGTTGCTACGACATGGAGGTGCCAGATCTCGGGCATC 2760
QY 2992 GACCTGCAAGTGCACATCGGCCCTGTATGGCAGCTTGGCTGGAGCTGGAGGTCAAGCAT 3051
DB 2761 GACCTGCAAGTGCACATCGGCCCTGTATGGCAGCTTGGCTGGAGCTGGAGGTCAAGCAT 2820
QY 3052 GGCCAGCTGGAGAACAGGCCCTTAACCTGCTCCACCGCGGCTCCACACTGCCGGA 3111
DB 2821 GGCCAGCTGGAGAACAGGCCCTTAACCTGCTCCACCGCGGCTCCACACTGCCGGA 2880
QY 3112 GCAGCCTTCTGCTCGGTGCAGATGTGCTCTGAAAAACAGGCTCAGCGGTTCCAGG 3171
DB 2881 GCAGCCTTCTGCTCGGTGCAGATGTGCTCTGAAAAACAGGCTCAGCGGTTCCAGG 2940
QY 3172 GGAATGCGAGCCCCCGGCTCAGTGGGAACAGGGCTTCGCAGCAGCAGGTTGGGGC 3231
DB 2941 GGAATGCGAGCCCCCGGCTCAGTGGGAACAGGGCTTCGCAGCAGCAGGTTGGGGC 3000
QY 3232 AAGCAGATGCTCCCTCAGGATTTACACTGAGCCCTGCCCAACCTGCTGAAAAAATCAT 3291
DB 3001 AAGCAGATGCTCCCTCAGGATTTACACTGAGCCCTGCCCAACCTGCTGAAAAAATCAT 3060
QY 3292 CGGCCACGCTGAAGACAGAGAGGAGTGGCAGGAGTTACCTGGGAAAAACAAACAGGGA 3351
DB 3061 CGGCCACGCTGAAGACAGAGAGGAGTGGCAGGAGTTACCTGGGAAAAACAAACAGGGA 3120
QY 3352 TCTTTTCTGCCCCCTGCTCCAGTCCAGTTGGCTCA 3387

Db	3121	TCTTTTTCGCCCTGCTCCAGTCGAGTGCGCCTGA	3156
RESULT 11	AR429673		
LOCUS	AR429673	2844 bp	DNA linear PAT 18-DEC-2003
DEFINITION	Sequence 2 from patent US 6645728.		
ACCESSION	AR429673		
VERSION	AR429673.1	GI:40189976	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2844)		
AUTHORS	Greene,W.C., Lin,X. and Gelezuinas,R.		
TITLE	Inhibitor of the inflammatory response induced by TNF.alpha. and IL-1		
JOURNAL	Patent: US 6645728-A 2 11-NOV-2003;		
FEATURES	Location/Qualifiers 1..2844 /organism="unknown" /mol_type="genomic DNA"		
ORIGIN			
Query Match	61.7%; Score 2837.6; DB 6; Length 2844;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 2840; Conservative	0; Mismatches 4; Indels 0; Gaps 0;		
Qy	232	ATGGCACTGATGGAATGGCTGCCTCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGNAG	291
Db	1	ATGGCAGTGATGGAATATGGCTGCTCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGNAG	60
Qy	292	GAACTCCCCAAGCCAAAAGGAGAAGACCCTGGCCACTGGGGGAAGAAACAGAGCTCCGTCTAC	351
Db	61	GAACTCCCCAAGCCAAAAGGAGAAGACCCTGGCCACTGGGGGAAGAAACAGAGCTCCGTCTAC	120
Qy	352	AAGCTTGAGGCGGTGGAGAAAGAGCCCTGTGTCTTCGCGAAAGTGGGAGATCTCTGAATGAC	411
Db	121	AAGCTTGAGGCGGTGGAGAAAGAGCCCTGTGTCTTCGCGAAAGTGGGAGATCTCTGAATGAC	180
Qy	412	GTGATTACCAAGGCGCAGCCNAGGAAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATC	471
Db	181	GTGATTACCAAGGCGCAGCCNAGGAAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATC	240
Qy	472	ATCGCCCAAGGCTGAGTGTGAGAATAGCCAAGAGTTTCAGCCCCCACCTTTTCAGAACGCATT	531
Db	241	ATCGCCCAAGGCTGAGTGTGAGAATAGCCAAGAGTTTCAGCCCCCACCTTTTCAGAACGCATT	300
Qy	532	TTCATCGCTGGGTCAAACAAGTAGCAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAT	591
Db	301	TTCATCGCTGGGTCAAACAAGTAGCAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAT	360
Qy	592	GTGSCCATGCTACAGAGGGCAAAATGGCCCGTGTGTGTTGGAAGGGAAGCGTCGAGC	651
Db	361	GTGSCCATGCTACAGAGGGCAAAATGGCCCGTGTGTGTTGGAAGGGAAGCGTCGAGC	420
Qy	652	AAAGCCCGGAAGAAACCGAAGAGAGAGCTCAAAGTCCCTGGCTCATGTCAGAGATGGCC	711
Db	421	AAAGCCCGGAAGAAACCGAAGAGAGAGCTCAAAGTCCCTGGCTCATGTCAGAGATGGCC	480
Qy	712	TTGGCCAAACCCCTCCCGAGGACCCCTGAGCAGGAGAGCTGCACCATCCCGAGTGCAGGAG	771
Db	481	TTGGCCAAACCCCTCCCGAGGACCCCTGAGCAGGAGAGCTGCACCATCCCGAGTGCAGGAG	540
Qy	772	GATGAGTCTCACTCGGCGGCCCONATGTTAGAAAAACACCCCGAGTTCACCAAGCCTCTG	831
Db	541	GATGAGTCTCACTCGGCGGCCCONATGTTAGAAAAACACCCCGAGTTCACCAAGCCTCTG	600
Qy	832	AAGGAACAGGCGCTTGGGCAACTCTGTTTTTAAGCAGCTTGGCGAGGCGCTACGGCCGGCT	891
Db	601	AAGGAACAGGCGCTTGGGCAACTCTGTTTTTAAGCAGCTTGGCGAGGCGCTACGGCCGGCT	660
Qy	892	CTGCGCTCGATCAGAACTCCCAAACTGATCAGCGCCCTTGCATAATGTCTGAACACACGTGTGG	951

Db 1741 AGCAGTGTGATGATGCTGCAATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTC 1800

QY 2032 CGAGGSCGCTCTGCTCAAGATTGCGACGAGCCTCGCCCTGTGAGGGAGATCCACCC 2091

Db 1801 CGAGGSCGCTCTGCTCAAGATTGCGACGAGCCTCGCCCTGTGAGGGAGATCCACCC 1860

QY 2092 TCTGTGCGCCCTCTCACAGCCAGGCCATCCAAAGAGGGCTGAGAAAGAGCCATCCAC 2151

Db 1861 TCTGTGCGCCCTCTCACAGCCAGGCCATCCAAAGAGGGCTGAGAAAGAGCCATCCAC 1920

QY 2152 CGCTGTCTGACGCGAGCTGGAGGGAAAGTGAACCGGGCACTACAGAACTGGAGGT 2211

Db 1921 CGCTGTCTGACGCGAGCTGGAGGGAAAGTGAACCGGGCACTACAGAACTGGAGGT 1980

QY 2212 CTGAAGAGCCCTTGAGGGGGAATATAAAGAACCAAGACATCCACCGCAATCAAGC 2271

Db 1981 CTGAAGAGCCCTTGAGGGGGAATATAAAGAACCAAGACATCCACCGCAATCAAGC 2040

QY 2272 AATTACCACAGACCTTCATGCCAGCGAGAGACTTTTCGCCAAGGGCCCCAGGGCCC 2331

Db 2041 AATTACCACAGACCTTCATGCCAGCGAGAGACTTTTCGCCAAGGGCCCCAGGGCCC 2100

QY 2332 CGGCAGCTGAGAGACAAAGCAGAGCCCTTAAGTCCAGCTCTCTCCCAACAGAG 2391

Db 2101 CGGCAGCTGAGAGACAAAGCAGAGCCCTTAAGTCCAGCTCTCTCCCAACAGAG 2160

QY 2392 CCCCAGAGCCAAAAGTCTCTCCCTTGAGCCAGCGCCCTTGAGCAAGGAGGTCTGGATGG 2451

Db 2161 CCCCAGAGCCAAAAGTCTCTCCCTTGAGCCAGCGCCCTTGAGCAAGGAGGTCTGGATGG 2220

QY 2452 GAACCCCTTACCTCTGTCTCCCTGGAGCAGCGCCCTTGCCAGAAACCCAGCTCACCAG 2511

Db 2221 GAACCCCTTACCTCTGTCTCCCTGGAGCAGCGCCCTTGCCAGAAACCCAGCTCACCAG 2280

QY 2512 CGGAAAGCAACCGTCCGAGCAGGAACTGACAGAGCTGGAAATAGAATATTCTCTAAC 2571

Db 2281 CGGAAAGCAACCGTCCGAGCAGGAACTGACAGAGCTGGAAATAGAATATTCTCTAAC 2340

QY 2572 AGCTGTCTCCAGCAATTTCTCTGAGGAGCAGGAGCAATTTCTCTCGTCCCTCAGATC 2631

Db 2341 AGCTGTCTCCAGCAATTTCTCTGAGGAGCAGGAGCAATTTCTCTCGTCCCTCAGATC 2400

QY 2632 GACAGCTCTCCCTGTGAGTACAGTGAAGAAACCCATCAAGGCCCTCTCAAGCTCG 2691

Db 2401 GACAGCTCTCCCTGTGAGTGAAGAAACCCATCAAGGCCCTCTCAAGCTCG 2460

QY 2692 CGGGAACCTGAGCTCAGGCGTACATCTCTGAGAGCAGCGCGAGGCTCGAAGCTCC 2751

Db 2461 CGGGAACCTGAGCTCAGGCGTACATCTCTGAGAGCAGCGCGAGGCTCGAAGCTCC 2520

QY 2752 AGCTGGACATGCTGTGCGCGGGGGCCCGACCGACCCCAAGCTATTCAATGGT 2811

Db 2521 AGCTGGACATGCTGTGCGCGGGGGCCCGACCGACCCCAAGCTATTCAATGGT 2580

QY 2812 GTGAAGTCCAAATACAGTCTTTAATGGTGAACACCTGCACATCCGGAGTTTCCACCG 2871

Db 2581 GTGAAGTCCAAATACAGTCTTTAATGGTGAACACCTGCACATCCGGAGTTTCCACCG 2640

QY 2872 GTCAAAGTGGAGACATCGCACTCGGCATCAGACGCGAGATCCCAAGCTTTCAGC 2931

Db 2641 GTCAAAGTGGAGACATCGCACTCGGCATCAGACGCGAGATCCCAAGCTTTCAGC 2700

QY 2932 TTGFTCACAAGAGCGGCGCTCTGCTACGACATGGAGTCCAGACTCGGGCATC 2991

Db 2701 TTGFTCACAAGAGCGGCGCTCTGCTACGACATGGAGTCCAGACTCGGGCATC 2760

QY 2992 GACCTGCAGTGACACTGCGCCCTGATGGAGCTTTCGCTGAGCTGGAGGTCAAGCAT 3051

Db 2761 GACCTGCAGTGACACTGCGCCCTGATGGAGCTTTCGCTGAGCTGGAGGTCAAGCAT 2820

QY 3052 GGCAGCTGGAGAAACAGGCCCTTAA 3075

Db 2821 GGCAGCTGGAGAAACAGGCCCTTAA 2844

RESULT 12
CQ845924
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
Yamamoto,O., Isono,Y., Nagai,K. and Irie,R.
Full-length human cdna
Patent: EP 1440981-A 4571 28-JUL-2004;
Research Association for Biotechnology (JP)
LOCATION/Qualifiers
FEATURES
1..3152
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 58.2%; Score 2673.4; DB 6; Length 3152;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 36; Indels 10; Gaps 7;
QY 1764 TCTGATCGGGACGTCAAAGCTGACAACTGCTCTCTGCCAGCATGGAGCCAGCCAGC 1823

Db 340 TCCC CGGGCGGCTGGCAGCTGACACGCTCTCTGCCAGCATGGAGCCAGCCAGC 399

QY 1824 CTTCTGTGACTTTGGCCATGTGTGTCTTTCAACTGATGCCCTCGGAAAGTCTTTGT 1883

Db 400 CTTCTGTGACTTTGGCCATGTGTGTCTTTCAACTGATGCCCTCGGAAAGTCTTTGT 459

QY 1884 CACAGGGACTACATCCCTGGCACAGACCCACATGCTCGGAGTGTGCTGGGCGAG 1943

Db 460 CACAGGGACTACATCCCTGGCACAGACCCACATGCTCGGAGTGTGCTGGGCGAG 519

QY 1944 GAGCTGCACGCCCAAGGTGGATGTCTGGAGCAGCTGTGTATGCTGCACATGCTCAA 2003

Db 520 GAGCTGCACGCCCAAGGTGGATGTCTGGAGCAGCTGTGTATGCTGCACATGCTCNA 579

QY 2004 CGGCTGCCACCCCTGGACTCAGTTCTTCGAGGGCGCTCTGCTCTCAAGATTGCCAGGA 2063

Db 580 CGGCTGCCACCCCTGGACTCAGTTCTTCGAGGGCGCTCTGCTCTCAAGATTGCCAGGA 639

QY 2064 GCCTCCGCTGTGAGGGAGATCCACCTCTGCGGCCCTCTCACAGCCCGGCCATCCA 2123

Db 640 GCCTCCGCTGTGAGGGAGATCCACCTCTGCGGCCCTCTCACAGCCCGGCCATCCA 699

QY 2124 AGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGACGGAGCTGGAGGGAAGGT 2183

Db 700 AGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGACGGAGCTGGAGGGAAGGT 759

QY 2184 GAA CGGGGCA CTACAGCAAGTGGAGGTCTGAAGAGCCCTTGGAGGGGAGAAATATAAGA 2243

Db 760 GAA CGGGGCA CTACAGCAAGTGGAGGTCTGAAGAGCCCTTGGAGGGGAGAAATATAAGA 819

QY 2244 ACCAAGACATCCCGCCCAATCAGCCAAATACACAGACCCCTCCATGCCAGCCGAG 2303

Db 820 ACCAAGACATCCCGCCCAATCAGCCAAATACACAGACCCCTCCATGCCAGCCGAG 879

QY 2304 AGAGCTTTGGCAAGGGCCCCAGGGCCCCGGCTGAGTGAAGAGCAACAGGCGAGCCGCC 2363

Db 880 AGAGCTTTGGCAAGGGCCCCAGGGCCCCGGCTGAGTGAAGAGCAACAGGCGAGCCGCC 939

QY 2364 TAAGCTCCAGCTCTCTCTCCCAACAGAGCCGCCAGAGCAACAAAGTCTCTCTCCCTTGAC 2423

Db 940 TAAGCTCCAGCTCTCTCTCTCCCAACAGAGCCGCCAGAGCAACAAAGTCTCTCTCCCTTGAC 999

QY 2424 TTTGAGCAAGGAGTCTGGGATGTGGGAACCCCTTACCTCTGTCCTCCCTGGAGCCAGC 2483
DB 1000 TTTGAGCAAGGAGTCTGGGATGTGGGAACCCCTTACCTCTGTCCTCCCTGGAGCCAGC 1059
QY 2484 CCCTGCCAGAAACCCAGCTCACAGAGCGGAAGCAACGTCCTGGAGCAGGAACCTGCA 2543
DB 1060 CCCTGCCAGAAACCCAGCTCACAGAGCGGAAGCAACGTCCTGGAGCAGGAACCTGCA 1119
QY 2544 GCAGCTGGAATAGAAATATTCTCAACAGAGCTGTCCAGGCAATTTCTCTGGAGGAGCA 2603
DB 1120 GCAGCTGGAATAGAAATATTCTCAACAGAGCTGTCCAGGCAATTTCTCTGGAGGAGCA 1179
QY 2604 GGAGCAAAATCTCTGTGCTCAGCATCGACAGCCTCTCCCTGTGCGATGACAGTGA 2663
DB 1180 GGAGCAAAATCTCTGTGCTCAGCATCGACAGCCTCTCCCTGTGCGATGACAGTGA 1239
QY 2664 GAAACCATCAAGGCTCTCAAGCTCGGGGACACCTGAGCTCAGGCTGACACTCTG 2723
DB 1240 GAAACCATCAAGGCTCTCAAGCTCGGGGACACCTGAGCTCAGGCTGACACTCTG 1299
QY 2724 GAGCAGCAGAGCGCTCGAAGCTCCAGCTGGAAATGATGTGTGGCCCGGGGCGGCC 2783
DB 1300 GAGCAGCAGAGCGCTCGAAGCTCCAGCTGGAAATGATGTGTGGCCCGGGGCGGCC 1359
QY 2784 CACCGACACCCCAAGCTATTCTCAATGTGTGAAAGTCCAAATACAGTCTCTTAATGTGA 2843
DB 1360 CACCGACACCCCAAGCTATTCTCAATGTGTGAAAGTCCAAATACAGTCTCTTAATGTGA 1419
QY 2844 ACACTGACATCCGGGATTCACCGGGTCAAGTGGGAGACATGCGCACTGCGATCAG 2903
DB 1420 ACACTGACATCCGGGATTCACCGGGTCAAGTGGGAGACATGCGCACTGCGATCAG 1479
QY 2904 CAGCAGATCCAGCTGAGCTTCAGCTTGGTCAACCAAGCGGCGAGCTGTCGCTA 2963
DB 1480 CAGCAGATCCAGCTGAGCTTCAGCTTGGTCAACCAAGCGGCGAGCTGTCGCTA 1539
QY 2964 CGACATGAGGTGCCAGACTCGGGATCGACCTGCAGTGCACACTGGGCCCTGATGGCAG 3023
DB 1540 CGACATGAGGTGCCAGACTCGGGATCGACCTGCAGTGCACACTGGGCCCTGATGGCAG 1599
QY 3024 CTTCCCTGGAGCTGGAGGTCAAGCATGGCCAGCTGGAGAACAGGCGCTTAACCTGCC 3083
DB 1600 CTTCCCTGGAGCTGGAGGTCAAGCATGGCCAGCTGGAGAACAGGCGCTTAACCTGCC 1659
QY 3084 TCCACCGCGCTCCACACTGCGGGAAGCAGCTTCTGCTCGGTGACAGCATGCTGCC 3143
DB 1660 TCCACCGCGCTCCACACTGCGGGAAGCAGCTTCTGCTCGGTGACAGCATGCTGCC 1718
QY 3144 TGAACACACAGCTCAGCGCTTCCAGGGGATTGCGAGCGCTTCCAGCTCAGTGGGAAC 3203
DB 1719 TG-AACACAGCTCAGCGCTTCCAGGGGATTGCGAGCGCTTCCAGCTCAGTGGGAC 1776
QY 3204 CAGGCTTCGAGCAGCAAGTGGGGCAAGCAAGTGCCTCCAGGATTTTACACCTGA 3263
DB 1777 CAGGCTTCGAGCAGCAAGTGGGGCAAGCAAGTGCCTCCAGGATTTTACACCTGA 1836
QY 3264 GCCCTGCCACCTCTGCTGAACAAACA-TCCGCGACGTGAAGAGACAGAGAGATGGC 3322
DB 1837 GCCCTGCCACCTCTGCTGAGAAACACTCCGCCACGTGAAGAGACAG-AGGAGGATGGC 1895
QY 3323 AGGAGTTACCTGGGGAACAAACAGGGATCTTTTCTGCCCTGCTCCAGTCCAGTTGG 3382
DB 1896 AGGAGTTACCTGGGGAACAAACAGG-----ATCTTCTGCTGCTCCAGTCCAGTTGG 1951
QY 3383 CTTGACCGCTTGGATCAGTCACTATTGTTGGCAGACAGGGGAGCAGCTTCCAGCT 3442
DB 1952 CTTGACCGCTTGGATCAGTCACTATTGTTGGCAGACAGGGGAGCAGCTTCCAGCT 2011
QY 3443 GGGTCAGAAAGGGTGGGCGAGCCCTTCGGCCCTCACCTCAGGCTGCTGTGAGATGT 3502
DB 2012 GGGTCAGAAAGGGTGGGCGAGCCCTTCGGCCCTCACCTCAGGCTGCTGTGAGATGT 2071

QY 3503 CAAGTGTGTAAAGGGCCCAAACTCAGGTTCACTGAGAACCAAGGTACAGAGTATGCCCGC 3562
DB 2072 CAAGTGTGTAAAGGGCCCAAACTCAGGTTCACTGAGAACCAAGGTACAGAGTATGCCCGC 2131
QY 3563 CCCTAGAGTTAAAGGGGGCCCTCTAAACCCCTTGCCTGGCCTCACCTGGCCAGCTCACCCCT 3622
DB 2132 CCCTAGAGTTAAAGGGGGCCCTCTAAACCCCTTGCCTGGCCTCACCTGGCCAGCTCACCCCT 2191
QY 3623 TTTGGGTGTAGGGGAAAGAAATGCTCAGCCCTGGGAAAGGCTCCCTGGTGTAGAAATACACC 3682
DB 2192 TTTGGGTGTAGGGGAAAGAAATGCTCAGCCCTGGGAAAGGCTCCCTGGTGTAGAAATACACC 2251
QY 3683 ACTTTTCAGGTTGTGTGCAACACACAGGTCTGAGTTGACCTCTGGTTCAGCCAGGACCAA 3742
DB 2252 ACTTTTCAGGTTGTGTGCAACACACAGGTCTGAGTTGACCTCTGGTTCAGCCAGGACCAA 2311
QY 3743 GAAGTGTGTAAAGTGAAGTGGTCTCAGTCCCCCAGACATGTGCCCTTTGCTGTGGCTA 3802
DB 2312 GAAGTGTGTAAAGTGAAGTGGTCTCAGTCCCCCAGACATGTGCCCTTTGCTGTGGCTA 2371
QY 3803 CCACCTTTCCCGAGCAGCAGCGCCCGAGCCCTTCAGGCCACAGCACTGCCCCAGACTC 3862
DB 2372 CCACCTTTCCCGAGCAGCAGCGCCCGAGCCCTTCAGGCCACAGCACTGCCCCAGACTC 2431
QY 3863 GCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGAGGATATGCTGTGACAGG 3922
DB 2432 GCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGAGGATATGCTGTGACAGG 2491
QY 3923 AACAGTCTGTGGATGGACATGATCAGTCTTAAGGAAAGCAGCAGAGAGAGAGG-TCCGGC 3981
DB 2492 AACAGTCTGTGGATGGACATGATCAGTCTTAAGGAAAGCAGCAGAGAGAGAGCGTCCGGC 2551
QY 3982 GCGCCAGCCCACTATCAGTGTCCAGCGTCTGGTTCGCCAGAGCAGCACTCAGCATCAC 4041
DB 2552 GCGCCAGCCCACTATCAGTGTCCAGCGTCTGGTTCGCCAGAGCAGCACTCAGCATCAC 2611
QY 4042 ACTGACACTCACCTGCCCTGCCCTGGCCAGAGGGTACTGCCAGCGGCACTTTGCACTC 4101
DB 2612 ACTGACACTCACCTGCCCTGCCCTGGCCAGAGGGTACTGCCAGCGGCACTTTGCACTC 2671
QY 4102 TGATGACCTCAAGCACTTTTCACTGGCTGCCCTTCGGCAGGGCAGGGCAGGCGAGTAC 4161
DB 2672 TGATGACCTCAAGCACTTTTCACTGGCTGCCCTTCGGCAGGGCAGGGCAGGCGAGTAC 2731
QY 4162 TGTAGGAGCAGTACCAAGCAGGAGATGGGTGAAGGAGCAGCTTTGAGCTGTCCACAT 4221
DB 2732 TGTAGGAGCAGTACCAAGCAGGAGATGGGTGAAGGAGCAGCTTTGAGCTGTCCACAT 2791
QY 4222 GCATGTGACTCCTCAAACTCTTCCAGATTTCTCTAAGAAATAGCACCCCTTCCCAATTG 4281
DB 2792 GCATGTGACTCCTCAAACTCTTCCAGATTTCTCTAAGAAATAGCACCCCTTCCCAATTG 2851
QY 4282 CCCAGCTTACCTCTTCTCCAGGGAGTCTCAGGACTCAGTACGATTAATACAGC 4341
DB 2852 CCCAGCTTACCTCTTCTCCAGGGAGTCTCAGGACTCAGTACGATTAATACAGC 2911
QY 4342 TGTCAATCGTCAGGGGTGTCTGCTAGCCTCAACCTCTCTGGGGCAGGGGAGCGCGAGACT 4401
DB 2912 TGTCAATCGTCAGGGGTGTCTGCTAGCCTCAACCTCTCTGGGGCAGGGGAGCGCGAGACT 2971
QY 4402 CCGTGGGAGAGCTCATTTCCACATCTTGGCAAGCAGCTTTTGTCCAGCTGTCCACATT 4461
DB 2972 CCGTGGGAGAGCTCATTTCCACATCTTGGCAAGCAGCTTTTGTCCAGCTGTCCACATT 3031
QY 4462 GAGTCAGACTGCTCCCGGGGAGAGAGCCCGGGCCCCCAGCAGCATTAAGAACTGCGAGCTT 4521
DB 3032 GAGTCAGACTGCTCCCGGGGAGAGAGCCCGGGCCCCCAGCAGCATTAAGAACTGCGAGCTT 3091
QY 4522 GGTACTCGAGTCTGGTGTGAGAGACTCTTTGTAGCAATAAAGTTTGGGGTGTGA 4581
DB 3092 GGTACTCGAGTCTGGTGTGAGAGACTCTTTGTAGCAATAAAGTTTGGGGTGTGA 3151
QY 4582 C 4582

QY 2904 CAGCCAGATCCAGCTGAGCCTTTCAGCTTGGTACCAAGACGGGAGCCTTTCGCTA 2963
Db 1480 CAGCCAGATCCAGCTGAGCCTTTCAGCTTGGTACCAAGACGGGAGCCTTTCGCTA 1539
QY 2964 CGACATGGAGGTGCGAGCTCGGCGATCGAGCTCGAGTGCACATCGGCCCTGATGCGAG 3023
Db 1540 CGACATGGAGGTGCGAGCTCGGCGATCGAGCTCGAGTGCACATCGGCCCTGATGCGAG 1599
QY 3024 CTTCCCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTTAACCCCTGCC 3083
Db 1600 CTTCCCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTTAACCCCTGCC 1659
QY 3084 TCCACGCCCGCTCCACATCGCCGGAAGAGCAGCTTCTCTGCTCGGTGCGATGCTGCC 3143
Db 1660 TCCACGCCCGCTCCACATCGCCGGAAGAGCAGCTTCTCTGCTCGGTGCGATGCTGCC 1718
QY 3144 TGAAGAACACAGGCTCAGCCGTTCCAGGGGATTCGACGCCCTCCGCTTCACAGTGGGAAC 3203
Db 1719 TG-AAGAACACAGGCTCAGCCGTTCCAGGGATTCGACGCCCTCCGCTTCACAGTGGGAAC 1776
QY 3204 CAGGCCCTCGCAGCAGCAAGGTGGGGCAAGCAGAAATGCTCCAGAGATTTTCACACCTGA 3263
Db 1777 CAGGCCCTCGCAGCAGCAAGGTGGGGCAAGCAGAAATGCTCCAGAGATTTTCACACCTGA 1836
QY 3264 GCCCTGCCCACTCTGTGAAAAAACA-TCCGCCACGTGAAGAGACAGAGAGATGGC 3322
Db 1837 GCCCTGCCCACTCTGTGAAAAAACA-TCCGCCACGTGAAGAGACAG-AGGAGGATGGC 1895
QY 3323 AGGAGTTACCTGGGGAAACAAACAGGATCTTTTCTGCCCTCTCCAGTCCAGTGG 3382
Db 1896 AGGAGTTACCTGGGGAAACAAACAGG-ATCTTCTCTGCCCTCTCCAGTCCAGTGG 1951
QY 3383 CCTGACCCGCTTGGATCAGTCAACATTTGTTGGCAGACAGGGGAGAGCAGCTTCCAGCCT 3442
Db 1952 CCTGACCCGCTTGGATCAGTCAACATTTGTTGGCAGACAGGGGAGAGCAGCTTCCAGCCT 2011
QY 3443 GGGTCAGAGGGGTGGGAGCCCTTCGGCCCTTCACCTCCAGGCTGTGTGAGAGTGT 3502
Db 2012 GGGTCAGAGGGGTGGGAGCCCTTCGGCCCTTCACCTCCAGGCTGTGTGAGAGTGT 2071
QY 3503 CAAGTGTGTAAGGGCCCAAACTCAGGTTCAAGTTCAGTGCAGAACAGGTCACAGGATGCCGC 3562
Db 2072 CAAGTGTGTAAGGGCCCAAACTCAGGTTCAAGTTCAGTGCAGAACAGGTCACAGGATGCCGC 2131
QY 3563 CCGTAGGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCCTCACTGGCCAGCTCACCCCT 3622
Db 2132 CCGTAGGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCCTCACTGGCCAGCTCACCCCT 2191
QY 3623 TTTGGGTGTAAGGGGAAAGATGCTGACCTGGGAAAGGCTCCCTGTGTGAGTAATACACAC 3682
Db 2192 TTTGGGTGTAAGGGGAAAGATGCTGACCTGGGAAAGGCTCCCTGTGTGAGTAATACACAC 2251
QY 3683 ACTTTTCAGGTTGTGCAACAGCTCCTGAGTTGACCTCTGTTTCAAGCCAGGACCAAA 3742
Db 2252 ACTTTTCAGGTTGTGCAACAGCTCCTGAGTTGACCTCTGTTTCAAGCCAGGACCAAA 2311
QY 3743 GAAGGTGTGAAGTGAAGTGTCTCAGTCCCAAGACATGTGCCCTTTGCTGTGGCTA 3802
Db 2312 GAAGGTGTGAAGTGAAGTGTCTCAGTCCCAAGACATGTGCCCTTTGCTGTGGCTA 2371
QY 3803 CCACCTTTCCAGACAGCAGGCCCCGAGCCCCCTTCAGGCCCAAGCAGCTGCCAGACTC 3862
Db 2372 CCACCTTTCCAGACAGCAGGCCCCGAGCCCCCTTCAGGCCCAAGCAGCTGCCAGACTC 2431
QY 3863 GCTGGCACTAGTTCCTCATCTGTAAAGGTGAAGGTGATGCAAGATATGCTTCAGAGG 3922
Db 2432 GCTGGCACTAGTTCCTCATCTGTAAAGGTGAAGGTGATGCAAGATATGCTTCAGAGG 2491
QY 3923 AACAGTCTGTGATGGACATCATCAGTCTGAAGAAAGCAGCAGACAGAGAGACG-TCCGGC 3981
Db 2492 AACAGTCTGTGATGGACATCATCAGTCTGAAGAAAGCAGCAGACAGAGAGACGCTCCGGC 2551
QY 3982 GCCCCAGCCCCCACTATCAGTGTCCAGGGTGTGTTCTCCCAAGACAGCAGCTCAGCATCAC 4041

Db 2552 GCCCCAGCCCCCACTATCAGTGTCCAGGGTGTGTTCTCCAGAGCAGCTCAGCATCAC 2611
QY 4042 ACTGACACTACCTCGCTGCCCTGCCAGAGGGTACTGCCAGCGCACTTTTGCATCTC 4101
Db 2612 ACTGACACTACCTCGCTGCCCTGCCAGAGGGTACTGCCAGCGCACTTTTGCATCTC 2671
QY 4102 TGATGACCTCAAGACACTTTTCATGCTGCCCTCTGGCAGGGCAGGGCAGGCGAGTGCAC 4161
Db 2672 TGATGACCTCAAGACACTTTTCATGCTGCCCTCTGGCAGGGCAGGGCAGGCGAGTGCAC 2731
QY 4162 TGATGAGCATAGCAAGCAGGAGATGGGTGAGAGGACACAGCTTTGAGCTGTCCACAT 4221
Db 2732 TGATGAGCATAGCAAGCAGGAGATGGGTGAGAGGACACAGCTTTGAGCTGTCCACAT 2791
QY 4222 GCATGTGACCTCTCAAACTCTTCCAGATTTCTTAAGAAATAGCACCCCTTCCCCATTG 4281
Db 2792 GCATGTGACCTCTCAAACTCTTCCAGATTTCTTAAGAAATAGCACCCCTTCCCCATTG 2851
QY 4282 CCCAGCTTAGCCTCTTCTCCAGGGAGTACTCAGGACTCAGTATGATTAATAATCAGC 4341
Db 2852 CCCAGCTTAGCCTCTTCTCCAGGGAGTACTCAGGACTCAGTATGATTAATAATCAGC 2911
QY 4342 TGTGAATCGTCAAGGGGTGTCTGCTAGCCTCAACCTCTTGGGGCAGGGGACGCCGAGCT 4401
Db 2912 TGTGAATCGTCAAGGGGTGTCTGCTAGCCTCAACCTCTTGGGGCAGGGGACGCCGAGCT 2971
QY 4402 CCGTGGGAGAGGCTCATTTCCACATCTTGGCAACAGACAGCCCTTTCAGCTGTCCACATT 4461
Db 2972 CCGTGGGAGAGGCTCATTTCCACATCTTGGCAACAGACAGCCCTTTCAGCTGTCCACATT 3031
QY 4462 GAGTCAGACTGTCTCCCGGAGAGAGCCCGGCCCCAGCACATAAAGAACTGAGCCTT 4521
Db 3032 GAGTCAGACTGTCTCCCGGAGAGAGCCCGGCCCCAGCACATAAAGAACTGAGCCTT 3091
QY 4522 GGTACTCAGAGCTGTGGTGTAGAGAACTCTTTGTAAGCAATAAGTTTGGGGTGTGA 4581
Db 3092 GGTACTCAGAGCTGTGGTGTAGAGAACTCTTTGTAAGCAATAAGTTTGGGGTGTGA 3151
QY 4582 C 4582
Db 3152 C 3152

RESULT 14

AF143094 2829 bp mRNA linear ROD 20-MAY-1999
LOCUS Mus musculus NF-kappaB inducing kinase (Nik) mRNA, complete cds.
DEFINITION AF143094
ACCESSION AF143094
VERSION AF143094.1 GI:4877962
KEYWORDS

SOURCE

Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 2829)
Suzuki, M., Kogishi, K., Serikawa, T. and Honjo, T.
Allymphoplasia is caused by a point mutation in the mouse gene

encoding Nf-kappa b-inducing kinase
Nat. Genet. 22 (1), 74-77 (1999)

JOURNAL

99251583
MEDLINE 10119865
PUBMED

REFERENCE

2 (bases 1 to 2829)
Shinkura, R., Kitada, K., Matsuda, F., Tashiro, K., Ikuta, K.,
Suzuki, M., Kogishi, K., Serikawa, T. and Honjo, T.

AUTHORS

Direct Submission
Submitted (14-APR-1999) Medical Chemistry, Kyoto University,
Yoshidakonoe-cho, Sakyo-ku, Kyoto 606-8315, Japan

JOURNAL

Location/Qualifiers
1. .2829
/organism="Mus musculus"
/mol_type="mRNA"

FEATURES

source

		/strain="C57BL/6" /db_xref="taxon:10090" /chromosome="11" map="between D11Mit331 and D11Mit360" 1. .2829 /gene="Nik" 1. .2829 /gene="Nik" /note="NIK" /codon_start=1 /product="NF-kappaB inducing kinase" /protein_id="AAD3152.1" /db_xref="GI:4877963" /translation="WAVMEVACPTPGSAVGQOKELAKAKEKTOSLGKQSCIFKLEA VKFSVPFCWKEILNDVITTKADKSGEGPPAISIIAQAEENSOFSTPFSERIFI AGSQOVSQGESUDQIPNNVAHATEGMARVCRGRKARKRKRKRSKSLKAGVA LAKPLRTPQESRDTQIPVQEDSPGLNLYARNVSQTKPLGGPGLGHLCPKQODEGLR PVLIPPELKLISPLQCLNHWKLHHPOATGPREHPHPFYSMPHPFFYFLEPWK FWMLSDALDKUAGVSGRPLPGLPHLSQLAHGDSQKPLPGPHLESSCPSRGALEKVP VVEZYIHALQGVSGRAHSLASAKTWSGSASLQRLGPETEDNEGVLLTEKLKPYD YBYREVHWHTQPRVGRGSFGEVHRMKDKQTFQCAVKVRLEFVFEVELVACGLS SPRIPLYCAVGEWPWNIEMELLEGLSLGOLIKOMGLPEDRALYVYLGQALRGLEYL HPRRIILHGDDKADNVLLSSDGSRAALCDFGHALCLQPDGLGKSLITGDIYIPGTETHWA PEVMVKGQADKVDYIWSRSCMWLMLNGCHFWTQIFRGLPCLIKIASFPFIRSIIPSC ALBITAQIQDEKURKEPVHSSAMELRRKVKALQEVGLKSPMKGEYKEPRPPIPOQA TCHQTLPTPRENPFAKNTADGAPPLPEPPEPSKAPALNSKEESGTWEPLPI SLSDPATKPPFPDRRATLPELELOQLEIELEFLNLSQPFSLEQOQILSCLISDLSL SLSDSEKNPSKASORDTLSSGVHWSNQAERTCSCSTALARGRPDTIPSYFNGV KVQIOSLNGEHLHIREFHRVKVDIATGISSQIPATAFSLVTKDQGPVCDYMEVPDSG IDLQCTLAPDGSFATWRVKHGLENNR"			
gene		Query Match 44.7%; Score 2054.6; DB 10; Length 2829; Best Local Similarity 83.5%; Pred. No. 0; Matches 2378; Conservative 0; Mismatches 444; Indels 27; Gaps 3;			
CDS					
		ORIGIN			
</					

Db 1621 TTGTGCTGCAACCTGACGGCCTAGGGAAATCCTTGTCTCA CAGGGGACTACATTCCTGGC 1680
Qy 1906 ACAGAGACCCACATGGCTCCGAGGTGTGTGGGCAGGAGCTGCGACGCCAAAGTGGAT 1965
Db 1681 ACGGAGACCCACATGACACAGAAAGTGTGTGATGGARAGCCCTGGATGCCAAGGTGGAC 1740
Qy 1966 GTCTGGAGCAGCTGTGTATGATGTGCATGTCTAAACGGGTGCGACCCCTGTGACTCAG 2025
Db 1741 ATCTGGAGCAGCTGTGTATGATGTCTCAATGTCTCAACGGGTGCGACCCCTGTGACTCAG 1800
Qy 2026 TTCTTTCCGAGGCGCTCTGCTCAAGATTGCCAGCAGCCTCCGCTGTGAGGGAGATC 2085
Db 1801 TACTTTCCAGGCCCGCTTTGTCTCAAGATTGCCAGCAGCCTCCACCGATCAGGGAGATC 1860
Qy 2086 CCACCTCTCTGCGCCCTCTCACAGCCAGGCCCTCAAGAGGGGCTGAGAAAGAGAGCC 2145
Db 1861 CCACCTCTCTGCGCACCCCTCACAGCCAGGCCCTCAAGAGGGGCTGAGAAAGAGCC 1920
Qy 2146 ATCCACCGCTGTCTGACGCGAGCTGGGAGGGAGGTGAACCGGGCACTACACAAAGTG 2205
Db 1921 GTCCACCGAGCATCTGCCATGAGCTTCGGAGGAAAGTGGCAAGGCCACTACAGAAAGTG 1980
Qy 2206 GGAGGTCTGAAGGCCCTTGAGGGGGAATATAAGAAACCAAGACATCCACCCGCAAT 2265
Db 1981 GGAGGTCTGAAGGCCCTTGAAAGGGAGAAATATAAGAAACCAAGACCTCCACCCCAAGAC 2040
Qy 2266 CAAGCCAAATACCAACAGACCTCTCATGCCAGCGAGAGAGCTTTCGCCAAGGGCCCCA 2325
Db 2041 CAAGCCACCTGCCACAGACCTACTACTCTCGCGGAGAGA----- 2081
Qy 2326 GGGCCCCGGCAGCTGAGGAGACAAAGCTCTCTCTGCTTGTGACGAGGAGGTCTGGG 2445
Db 2082 --GAACCCACCAAGCCAAAGCCCAACACAGACGGGGCTCTGAGCCTCAGCCTCTCTACCG 2139
Qy 2386 CCAGAGCCCCCAGAGGCCAAACAGTCTCTCTCTGCTTGTGACGAGGAGGTCTGGG 2445
Db 2140 CCAGAACCCACAGAACCGAGCAAGAGCGCCAGCCCTGAACTTGAGCAAGGAGGAGTCTGGC 2199
Qy 2446 ATGTGGGAACCTTACCTCTGCTCTCTCTGCTGAGCGCAGCCCTGCGAGAAACCCAGCTCA 2505
Db 2200 ACATGGGNAACCTTGCTCTCTCTCTGCTGAGCCAGCCACTGCGAAAGGCCCGAGCTTC 2259
Qy 2506 CCAGAGGGGAAGCAACCGTCCCGAGCAGAACTGACGAGCTGGAATAAGAAATATTTC 2565
Db 2260 CCAGACCGAGGGCAACCTTGCCAGAGCTGAGCTACAGCAACTGGAGATAGAACTGTTT 2319
Qy 2566 CTCACAGCCTGTCCAGCCATTTTCTCTGAGAGCAGGAGCAAAATTTCTCTGTGCTC 2625
Db 2320 CTCACAGCCTGTCCAGCCGTTCTCTCTGAGGAAACAGGAACAAATCTCTCTGCTGCTC 2379
Qy 2626 AGCATCGACAGCTCTCCCTGTGCGATGACAGTGAAGAAGACCCATCAAGGCCCTCTCAA 2685
Db 2380 AGCATCGACAGCTCTCTGCTGAGATGACAGTGAAGAAGATTCATCGAAGCCCTCTCAG 2439
Qy 2686 AGCTCGGGGACACCTTGAGCTCAGGGGTACACTCTCTGAGCAGCCAGCGCGAGGCTGA 2745
Db 2440 AGCTCAGGGACACCTTGAGTCTGGGTGCTGCTTGTGAACAGCAAGCTGAGGCAAGA 2499
Qy 2746 AGCTCAGCTGGAACATGTGCTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTC 2805
Db 2500 ACCTGCACTGCAAGCAGCGCTGGCGGGGGGGGGGCTACTGACATCCCGAGCTACTTC 2559
Qy 2806 AATGCTGGAAGTCCAAATACAGTCTCTTAATGTGAACACCTGACATCCGGGAGTTC 2865
Db 2560 AACGGGTGCAAGGTCCAGATCCAGTCTCTCAATGGCGAACHCTTGCATATCCGGGAATTC 2619
Qy 2866 CACCGGTCAAAGTGGGAGACATGCCACTGGCATCAGCAGCCAGATCCCAAGCTGCGAGCC 2925
Db 2620 CACCGGTCAAAGTGGGAGACATTTGCCACCGCATCAGCAGCCAGATCCCAAGCTGCGAGCT 2679
Qy 2926 TTCAGCTGTGTCACAAAGACGGGAGCTGTTTGTGCTACGACATGAGGAGTCCAGACTCG 2985
Db 2680 TTCAGCTGTGTCACAAAGATGAGCAGCTGTTTGTGCTATGACATGAGGAGTCCAGACTCG 2739

Qy 2986 GGCAATCGACCTGCAGTGCACACTGGCCCTGATGGCAGCTTCGCTTGAGCTGGAGGGTC 3045
Db 2740 GGCAATCGACCTGCAGTGCACACTGGCCCTGATGGCAGCTTCGCTTGAGCTGGAGGGTC 2799
Qy 3046 AAGCATGCCAGCTGGAGAAACAGGCCCTA 3074
Db 2800 AAGCATGGTCAGCTGGAGAACCCGACCTA 2828
RESULT 15
A66647
LOCUS A66647 2631 bp DNA linear PAT 29-MAR-1999
DEFINITION Sequence 3 from Patent WO9737016.
ACCESSION A66647
VERSION A66647.1 GI:4538139
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2631)
AUTHORS Wallach, D., Malinin, N., Boldin, M., Kovalenko, A. and Mett, I.
TITLE MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR
PREPARATION AND USE
JOURNAL Patent: WO 9737016-A 3 09-OCT-1997;
YEDA RES & DEV (IL)
COMMENT Other publication AU 2175597 19971022.
FEATURES
Location/Qualifiers
source 1..2631
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

Query Match 40.3%; Score 1851.2; DB 6; Length 2631;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2335; Conservative 4; Mismatches 159; Indels 81; Gaps 34;
Qy 2099 CCCCTCTCACAGCCAGCCAGGCATCCAGAGGGGCTGAGGAAAGAGCCATCCACCGCGTGT 2158
Db 1 CCCCTCTCACAGCCAGCCAGGCATCCAGAGGGGCTGAGGAAAGAGCCATCCACCGCGTGT 60
Qy 2159 CTGACGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 2218
Db 61 CTGACGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120
Qy 2219 GCCCTTGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 2278
Db 121 GCCCTTGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 180
Qy 2279 ACCAGACCTCCATGCCCGCAGCGAGAGCTTTTCGCAAGGGCCCGAGGGCCCGGCCAG 2338
Db 181 ACCAGACCTCCATGCCCGCAGCGAGAGCTTTTCGCAAGGGCCCGAGGGCCCGGCCAG 240
Qy 2339 CTGAGGAGACAAACAGCGAGAGCCCTTAAGCTTCCAGCCTCTCTCTCCACAGAGCCCGCAG 2398
Db 241 CTGAGGAGACAAACAGCGAGAGCCCTTAAGCTTCCAGCCTCTCTCTCCACAGAGCCCGCAG 300
Qy 2399 AGCCAAACAAAGTCTCTCTCTTGTGAGCAAGGAGGAGTCTGGGATGAGGAAACCT 2458
Db 301 AGCCAAACAAAGTCTCTCTCTTGTGAGCAAGGAGGAGTCTGGGATGAGGAAACCT 360
Qy 2459 TACCTGTCTCTCTCTGAGCCAGCCCTTCGCAAGAACCCAGCTCCACAGAGCGGAAG 2518
Db 361 TACCTGTCTCTCTCTGAGCCAGCCCTTCGCAAGAACCCAGCTCCACAGAGCGGAAG 420
Qy 2519 CAACCGTCCCGGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATTATTCCTCAACAGCCTGT 2578
Db 421 CAACCGTCCCGGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATTATTCCTCAACAGCCTGT 480
Qy 2579 CCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTGCTACGATCGACAGCC 2638
Db 481 CCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTGCTACGATCGACAGCC 540

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 22:58:14 ; Search time 1562.09 Seconds
(without alignments)
17417.120 Million cell updates/sec

Title: US-09-155-676B-6
Perfect score: 4596
Sequence: 1 agcgggggactgtgcgtg.....gatgacaaatgttaaaaaa 4596

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4596	100.0	4596	2	AAV03326
2	4578.4	99.6	4596	5	AAS79490
3	4578.4	99.6	4596	6	ABK51171
4	4578.4	99.6	4596	10	ADP85075
5	4578.4	99.6	4596	13	ADR25114
6	4440	96.6	4684	11	ACN44439
7	3149.6	68.5	3156	2	AAV73917
8	3149.6	68.5	3156	2	AAV71603
9	3149.6	68.5	3156	2	AAV69285
10	2839.2	61.8	2844	2	AAH87842
11	2673.4	58.2	3152	12	ADQ67410
12	2042.6	44.4	2828	11	ACN44437
13	1851.2	40.3	2631	2	AAV03324
14	1556.2	32.9	73967	11	ACN44438
15	1474.4	32.1	2760	5	AAS79491
16	1304	28.4	1385	4	AAH99262
17	1241	27.0	2070	12	ADP28582
18	627.6	13.7	722	5	AAS79492
19	480.2	10.4	42566	11	ACN44436
20	476	10.4	476	10	ADF79802

C	21	349.4	7.6	2475	11	ADM03069	Adm03069 Human CDN
	22	299.2	6.5	362	3	AAC02585	Aac02585 Human sec
	23	291.6	6.3	6673	6	ABQ67111	Abq67111 Human ang
	24	267.4	5.8	499	9	ACH15373	Ach15373 Human adu
C	25	243.6	5.3	6673	6	ABQ67112	Abq67112 Human ang
	26	171.8	3.7	500	12	ACH75442	Ach75442 Human gen
	27	166.8	3.6	293	12	ACH89170	Ach89170 Human gen
	28	89.8	2.0	707	12	ADH56283	Adh56283 Human MAP
	29	88.8	1.9	888	12	ADH56286	Adh56286 Human MAP
	30	88.8	1.9	900	12	ADH56282	Adh56282 Human MAP
	31	88.8	1.9	973	12	ADH56281	Adh56281 Human MAP
	32	88.8	1.9	1971	12	ADH56272	Adh56272 Human MAP
	33	88.8	1.9	2032	12	ADN61533	Adn61533 Human KPP
	34	88.8	1.9	2154	12	ADH56270	Adh56270 Human MAP
	35	88.8	1.9	2348	6	AAH34715	Aah34715 Human MEK
	36	88.8	1.9	2348	12	ADN71931	Adn71931 Human kin
	37	88.8	1.9	2488	13	ADR39801	Adr39801 Human ORF
	38	88.8	1.9	3150	12	ADI26096	Adi26096 Human CDN
	39	88.8	1.9	3286	3	AAC76868	Aac76868 Human ORF
	40	88.8	1.9	3295	13	ACN40256	Acn40256 Tumour-as
	41	88.8	1.9	3621	12	ADN71923	Adn71923 Human DKF
C	42	88.4	1.9	455	5	ABV22061	Abv22061 Human pro
C	43	88.4	1.9	455	5	ABV27896	Abv27896 Human pro
C	44	88.4	1.9	476	5	ABV12334	Abv12334 Human pro
C	45	88.4	1.9	504	5	ABV33478	Abv33478 Human pro

ALIGNMENTS

RESULT 1
AAV03326
ID AAV03326 standard; DNA; 4596 BP.
XX AC AAV03326;
XX AC AAV03326;
DT 15-APR-1998 (first entry)
XX DB DNA sequence of NF-kappaB inducing Kinase.
XX KW Human tumour necrosis factor receptor-associated factor 2; TRAF2;
KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;
KW intracellular signalling activity; acute hepatitis;
KW autoimmune-induced cell death; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 232..3075
FT FT /*tag= a
XX PN WO9737016-A1.
XX PD 09-OCT-1997.
XX PF 01-APR-1997; 97WO-IL000117.
XX PR 02-APR-1996; 96IL-00117800.
XX PR 26-AUG-1996; 96IL-00119133.
XX PA (YEDA) YEDA RES & DEV CO LTD.
XX PI Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;
XX WPI; 1997-503101/46.
XX DR P-PSDB; AAW42402.
XX PT DNA encoding tumour necrosis factor receptor-associated factor binding
XX PT molecule - used for modulation or mediation in cells of the activity of
XX NF-kB.
XX PS Claim 11; Fig 6; 127pp; English.
XX

CC The present sequence encodes a NF-kappa inducing kinase (NIK). The full
CC length sequence was obtained by PCR using clone 10 (RAV03324). NIK was
CC found to induce NF-kappas even more effectively than TRAF2. Proteins
CC capable of binding to TRAF2 and NIK were identified. The NIK or TRAF-2
CC binding proteins can be used for modulation or mediation in cells of NF-
CC kappaB activity or any other intracellular signalling activity modulated
CC or mediated by TRAF2. TRAF-binding proteins are especially used for
CC prevention or treatment of pathological conditions associated with NF-
CC kappaB induction, e.g. acute hepatitis, autoimmune-associated cell death,
CC e.g. death of the beta Langerhans cells or the pancreas that results in
CC diabetes, the death of cells in graft rejection, the death of
CC oligodendrocytes in the brain in multiple sclerosis, and AIDS-inhibited T
CC cell suicide which causes proliferation of the AIDS virus and hence the
CC AIDS disease. The proteins are also useful for screening of ligands
CC capable of binding to a protein, which are useful for modulating cellular
CC activity modulated/mediated by TRAF2
XX
SQ Sequence 4596 BP; 1055 A; 1402 C; 1298 G; 841 T; 0 U; 0 Other;

Query Match 100.0%; Score 4596; DB 2; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGGGGGGACGTGCGGTGGAACTGTAGCTGTGAAGGTGACTCTGTATACATTG 60
DB 1 ACGGGGGGACGTGCGGTGGAACTGTAGCTGTGAAGGTGACTCTGTATACATTG 60

QY 61 AGGATTTTGGAGGATGATATGTGTGGCAGAGGCACATAAACAGGACAGACCCCTT 120
DB 61 AGGATTTTGGAGGATGATATGTGTGGCAGAGGCACATAAACAGGACAGACCCCTT 120

QY 121 GCCCTGCTTTCTCCCAACCCAGGCTGACCTGTGTTCTCCAGGTCTGGATTCTA 180
DB 121 GCCCTGCTTTCTCCCAACCCAGGCTGACCTGTGTTCTCCAGGTCTGGATTCTA 180

QY 181 AGTGACTGCTCTGTGTGTGTGTCTCTCAGGATGAGCAACAAGCTGGAGATGGCAGTG 240
DB 181 AGTGACTGCTCTGTGTGTGTGTCTCTCAGGATGAGCAACAAGCTGGAGATGGCAGTG 240

QY 241 ATGGAAATGGCTGCGCAGGTGCCCCCTGGCTAGCAGTGGGGCAGCAGAACTCCCC 300
DB 241 ATGGAAATGGCTGCGCAGGTGCCCCCTGGCTAGCAGTGGGGCAGCAGAACTCCCC 300

QY 301 AAGCAAAGAGAGAGACGGCCACCTGGGGAAGAAACAGAGCTCCGTCTACAAGCTTGAG 360
DB 301 AAGCAAAGAGAGAGACGGCCACCTGGGGAAGAAACAGAGCTCCGTCTACAAGCTTGAG 360

QY 361 GCCGTGGAGAGAGCCCTGTGTCTCGGAAAGTGGGAGATCCTGAATGACGTGATTACC 420
DB 361 GCCGTGGAGAGAGCCCTGTGTCTCGGAAAGTGGGAGATCCTGAATGACGTGATTACC 420

QY 421 AAGGGCAGACCAAGGAGGCTCGAGGAGGCGCCAGCTGCCATCTCTATCATCGCCAG 480
DB 421 AAGGGCAGACCAAGGAGGCTCGAGGAGGCGCCAGCTGCCATCTCTATCATCGCCAG 480

QY 481 GCTGAGTGTGAGATAGCCAAAGATTCAGCCCCACCTTTTCAGAAACGCAATTTTCATCGCT 540
DB 481 GCTGAGTGTGAGATAGCCAAAGATTCAGCCCCACCTTTTCAGAAACGCAATTTTCATCGCT 540

QY 541 GGGTCCAAAACAGTACAGCCAGTCCGAGAGTCTTTGATCAGATCCCCAACATGTGGCCCAT 600
DB 541 GGGTCCAAAACAGTACAGCCAGTCCGAGAGTCTTTGATCAGATCCCCAACATGTGGCCCAT 600

QY 601 GCTACAGAGGCAAAATGGCCCGTGTGTGGAGGGAAGCGTCGCAAGAACCCCG 660
DB 601 GCTACAGAGGCAAAATGGCCCGTGTGTGGAGGGAAGCGTCGCAAGAACCCCG 660

QY 661 AAGAAACGGAAGAGAGAGCTCAAAAGTCCCTGGCTCATGAGAGTGGCCCTTTGGCCAAA 720
DB 661 AAGAAACGGAAGAGAGAGCTCAAAAGTCCCTGGCTCATGAGAGTGGCCCTTTGGCCAAA 720

QY 721 CCCCTCCCCAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGCAGGAGGATGAGTCT 780
DB 721 CCCCTCCCCAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGCAGGAGGATGAGTCT 780

DB 721 CCCCTCCCCAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGCAGGAGGATGAGTCT 780
QY 781 CCACCTGGCGCCCCATATGTTAGAAACACCCCCAGTTTCAACCAAGCTCTTGAAGGACCA 840
DB 781 CCACCTGGCGCCCCATATGTTAGAAACACCCCCAGTTTCAACCAAGCTCTTGAAGGACCA 840
QY 841 GGCCTTGGGCAACTCTGTTTTAAGCAGCTTGGCGAGGGCTTACGGCCGGCTCTGCTCTCGA 900
DB 841 GGCCTTGGGCAACTCTGTTTTAAGCAGCTTGGCGAGGGCTTACGGCCGGCTCTGCTCTCGA 900
QY 901 TCAGAACTCCACAAACTGATCAGCCCTTGCATATGTTGAACACACCTGTGGAAGTCTGCAC 960
DB 901 TCAGAACTCCACAAACTGATCAGCCCTTGCATATGTTGAACACACCTGTGGAAGTCTGCAC 960
QY 961 CACCCACAGAGGAGGAGCCCTGCTGCTGCCACGACACCCCTTCCCTCTATAGCAGACTG 1020
DB 961 CACCCACAGAGGAGGAGCCCTGCTGCTGCCACGACACCCCTTCCCTCTATAGCAGACTG 1020
QY 1021 CCTCATCCCTTCCCATTTCCACCCCTTCCAGCCCTTCCAGCCCTTGGAAACCTCACCCCTCTGGAGTCTTTC 1080
DB 1021 CCTCATCCCTTCCCATTTCCACCCCTTCCAGCCCTTGGAAACCTCACCCCTCTGGAGTCTTTC 1080
QY 1081 CTGGGCAAACTGGCCCTGTGTAGACAGCAGCAAAACCTTGGCTGACCCACACCTCTGAGCAAA 1140
DB 1081 CTGGGCAAACTGGCCCTGTGTAGACAGCAGCAAAACCTTGGCTGACCCACACCTCTGAGCAAA 1140
QY 1141 CTGGCCCTGTGTAGACAGTCCAAAGCCCTGCTGCTGGCCACACCTTGGAGCCCGAGTGCCTG 1200
DB 1141 CTGGCCCTGTGTAGACAGTCCAAAGCCCTGCTGCTGGCCACACCTTGGAGCCCGAGTGCCTG 1200
QY 1201 TCTCGTGTGCCATAGAGAAAGTTTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGC 1260
DB 1201 TCTCGTGTGCCATAGAGAAAGTTTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGC 1260
QY 1261 AGCTGTAGCTCAAGCCAGGCCCCACAGCTGACAGCTGGGCAAGAGCTGGGCGACGACGG 1320
DB 1261 AGCTGTAGCTCAAGCCAGGCCCCACAGCTGACAGCTGGGCAAGAGCTGGGCGACGACGG 1320
QY 1321 GGCTCCAGATCCCCGGGAGCCCGAGCCCAAAAACCTGAGGACAAACGAGGGTGTCTGCTCACT 1380
DB 1321 GGCTCCAGATCCCCGGGAGCCCGAGCCCAAAAACCTGAGGACAAACGAGGGTGTCTGCTCACT 1380
QY 1381 GAGAACTCAAGCCAGTGGATTTAGTATCCGAGAGAGTCCACTGGGCCACACCCAG 1440
DB 1381 GAGAACTCAAGCCAGTGGATTTAGTATCCGAGAGAGTCCACTGGGCCACACCCAG 1440
QY 1441 CTCCGCTGGGCGAGAGGCTCTTCCGAGAGGTGTCACAGATGGAGCAAGCAGACTGCG 1500
DB 1441 CTCCGCTGGGCGAGAGGCTCTTCCGAGAGGTGTCACAGATGGAGCAAGCAGACTGCG 1500
QY 1501 TTCCAGTGGCTGTCAAAAAGTGGCGCTGGAAGTATTTTCGGGCGAGAGGATGATGGCA 1560
DB 1501 TTCCAGTGGCTGTCAAAAAGTGGCGCTGGAAGTATTTTCGGGCGAGAGGATGATGGCA 1560
QY 1561 TGTGAGGATTTGACCTCAACCCAGATTTGCTCCCTTTGATGAGGCTGTGAGAGAGGGCT 1620
DB 1561 TGTGAGGATTTGACCTCAACCCAGATTTGCTCCCTTTGATGAGGCTGTGAGAGAGGGCT 1620
QY 1621 TGGGTCAACATCTTCAATGAGCTGTGGAAGTGGCTCCCTGGGCCAGCTGTGTCAGAGG 1680
DB 1621 TGGGTCAACATCTTCAATGAGCTGTGGAAGTGGCTCCCTGGGCCAGCTGTGTCAGAGG 1680
QY 1681 CAGGGCTGTCTCCAGAGACCGGGCCCTGTACTACTCTGGGCGAGGCCCTGGAGGGTCTG 1740
DB 1681 CAGGGCTGTCTCCAGAGACCGGGCCCTGTACTACTCTGGGCGAGGCCCTGGAGGGTCTG 1740
QY 1741 GAATACCTCACTCAGGAGGATTTGATGAGGAGCTCAAAAGCTGACAACTGTCTCTCTG 1800
DB 1741 GAATACCTCACTCAGGAGGATTTGATGAGGAGCTCAAAAGCTGACAACTGTCTCTCTG 1800
QY 1801 TCCAGGAGTGGAGGACCGAGCCCTCTGTGATCTTTGGCCATCTGTGTGTCTTCAACT 1860
DB 1801 TCCAGGAGTGGAGGACCGAGCCCTCTGTGATCTTTGGCCATCTGTGTGTCTTCAACT 1860

Db 4021 CAGAGCAGAGCTCAGCATCACACTGACACTCACCTGCTGCCCTGSCCAGAGGGTAC 4080
QY 4081 TGCAGCGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTGGCAG 4140
Db 4081 TGCAGCGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTGGCAG 4140
QY 4141 GGCAGGGCAGGCGAGTGACACTGTAGGAGCATGACGAGCCAGGATGGGGTGAAGGGAC 4200
Db 4141 GGCAGGGCAGGCGAGTGACACTGTAGGAGCATGACGAGCCAGGATGGGGTGAAGGGAC 4200
QY 4201 ACAGTCTTGAGCTGTCCACATGCATGTGACTCTCTCAAACTCTTCCAGATTTCTCTAAGA 4260
Db 4201 ACAGTCTTGAGCTGTCCACATGCATGTGACTCTCTCAAACTCTTCCAGATTTCTCTAAGA 4260
QY 4261 ATAGACACCCCTTCCCATTCGCCAGCTTAGCTCTTCCAGGGGAGCTACTCAGGA 4320
Db 4261 ATAGACACCCCTTCCCATTCGCCAGCTTAGCTCTTCCAGGGGAGCTACTCAGGA 4320
QY 4321 CTCAGTAGCATTAATCAGCTGTGAATCGTCAGGGGGTGTCTGTAGCCTCAACTCTCT 4380
Db 4321 CTCAGTAGCATTAATCAGCTGTGAATCGTCAGGGGGTGTCTGTAGCCTCAACTCTCT 4380
QY 4381 GGGCAGGGGACGCCGAGACTCCGTGGGAGAGCTCAITCCACATCTTGCCACAGCAGC 4440
Db 4381 GGGCAGGGGACGCCGAGACTCCGTGGGAGAGCTCAITCCACATCTTGCCACAGCAGC 4440
QY 4441 CTTTGTCCAGCTGTCCACATTCAGTCAGCTCTCCGGGGAGAGCCCGGCCCCCGAG 4500
Db 4441 CTTTGTCCAGCTGTCCACATTCAGTCAGCTCTCCGGGGAGAGCCCGGCCCCCGAG 4500
QY 4501 CACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTGTAA 4560
Db 4501 CACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTGTAA 4560
QY 4561 CAATAAAGTTGGGGTGTAGCAAAATGTTAAAAAAA 4596
Db 4561 CAATAAAGTTGGGGTGTAGCAAAATGTTAAAAAAA 4596

RESULT 2
AAS79490
ID AAS79490 standard; cDNA; 4596 BP.
XX AAS79490;
AC AAS79490;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #15294.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG15303.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 15294; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 4596 BP; 1055 A; 1401 C; 1299 G; 841 T; 0 U; 0 Other;

Query Match 99.6%; Score 4578.4; DB 5; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 5 GGGGACTGTGCGGTGTGGAACGTGTAGCTGTGTA-AGGTGACTCTGTACCATTGAGG 63
Db 5 GGGGACTGTGCGGTGTGGAACGTGTAGCTGTGTAAGGTGACTCTGTACCATTGAGG 64
QY 64 ATGTTTGAGGATGAGTATGTGGCAGGACACATAAACAGGACAGACCTTTGGC 123
Db 65 ATGTTTGAGGATGAGTATGTGGCAGGACACATAAACAGGACAGACCTTTGGC 124
QY 124 CCGCTTTTCTCCCCAACCCAGGCTGACCTGTGTTCTTCCCAGGTCTGGGATTTAAAGT 183
Db 125 CCGCTTTTCTCCCCAACCCAGGCTGACCTGTGTTCTTCCCAGGTCTGGGATTTAAAGT 184
QY 184 GACCTGCTCTGTGTTGTTGTTCTCTCAGGATGAGCACAAGCCTGGAGATGGCAGTGATG 243
Db 185 GACCTGCTCTGTGTTGTTGTTCTCTCAGGATGAGCACAAGCCTGGAGATGGCAGTGATG 244
QY 244 GAAATGGCTGCGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAAGAACTCCCCAAG 303
Db 245 GAAATGGCTGCGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAAGAACTCCCCAAG 304
QY 304 CCAAAAGGAGAAGACGCCGCCACTGCGGGAAGAAACAGAGCTCCGTTACAAAGTTGAGGCC 363
Db 305 CCAAAAGGAGAAGACGCCGCCACTGCGGGAAGAAACAGAGCTCCGTTACAAAGTTGAGGCC 364
QY 364 GTGGAGAAAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCTGAATGACGTGATTACCAAG 423
Db 365 GTGGAGAAAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCTGAATGACGTGATTACCAAG 424
QY 424 GGCACAGCCCAAGGAGGCTCCGAGGCGAGGCGAGCTGCTCATCTCATATCCCGAGGCT 483
Db 425 GGCACAGCCCAAGGAGGCTCCGAGGCGAGGCGAGCTGCTCATCTCATATCCCGAGGCT 484
QY 484 GAGTGTGAGAAATAGCCAAAGAGTTTCCAGCCCCACCTTTTTCAGAACGCAATTTTCATCGTGGG 543
Db 485 GAGTGTGAGAAATAGCCAAAGAGTTTCCAGCCCCACCTTTTTCAGAACGCAATTTTCATCGTGGG 544
QY 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATTCAGATCCCCCAACAATGTGCCCCATGCT 603
Db 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATTCAGATCCCCCAACAATGTGCCCCATGCT 604

Qy	604	ACAGGGCAAAATGGCCCGTGTGTGTTGGAAAGGAAAGCGTCGACGAAAGCCCGGAAG	663
Db	605	ACAGAGGCAAAATGGCCCGTGTGTGTTGGAAAGGAAAGCGTCGACGAAAGCCCGGAAG	664
Qy	664	AAACGGAAGAAGAGAGCTCAAGTCCCTGCTCATGACGAGTGGCCTTGCCCAAAACC	723
Db	665	AAACGGAAGAAGAGAGCTCAAGTCCCTGCTCATGACGAGTGGCCTTGCCCAAAACC	724
Qy	724	CTCCCCAGGACCCCTGACGAGAGAGTGCACCATCCAGTGCAGGAGATGAGTCTCCA	783
Db	725	CTCCCCAGGACCCCTGACGAGAGAGTGCACCATCCAGTGCAGGAGATGAGTCTCCA	784
Qy	784	CTCGCGCCCATATGTTAGAAAACCCCGCAGATTACCAAGCCTCTGAAGAAACAGGC	843
Db	785	CTCGCGCCCATATGTTAGAAAACCCCGCAGATTACCAAGCCTCTGAAGAAACAGGC	844
Qy	844	CTTGGCAACTCTGTTTTAAGCAGCTTGGCAGAGGCTTACGCGCGCTCTCCCTCGATCA	903
Db	845	CTTGGCNACTCTGTTTTAAGCAGCTTGGCAGAGGCTTACGCGCGCTCTCCCTCGATCA	904
Qy	904	GAACTCCACAAACTGATCAGCGCCCTTGCAATGTCTGAACCACTGTGGAAACTGCACCAC	963
Db	905	GAACTCCACAAACTGATCAGCGCCCTTGCAATGTCTGAACCACTGTGGAAACTGCACCAC	964
Qy	964	CCCCAGGAGGAGGCCCCCTGCGCTGCCCCAGCACCCCTTCCCTATAGCAGACTGCT	1023
Db	965	CCCCAGGAGGAGGCCCCCTGCGCTGCCCCAGCACCCCTTCCCTATAGCAGACTGCT	1024
Qy	1024	CATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAACCTCACTCTGGAGTCTTCTCTG	1083
Db	1025	CATCCCTTCCCATTTCCACCTCTCTAGCCCTTGGAAACCTCACTCTGGAGTCTTCTCTG	1084
Qy	1084	GGCAAACTGGCTGTGTAGACAGCAGAAACCTTTCCTGACCCACACTGAGCAAACTG	1143
Db	1085	GGCAAACTGGCTGTGTAGACAGCAGAAACCTTTCCTGACCCACACTGAGCAAACTG	1144
Qy	1144	GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCCACACTGGAGCCAGTGCTGTCT	1203
Db	1145	GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCCACACTGGAGCCAGTGCTGTCT	1204
Qy	1204	CGTGGTCCCATGAGAGTTTCTGTGGAGAAATACCTAGTGCATGCTCTGCAAGGAGC	1263
Db	1205	CGTGGTCCCATGAGAGTTTCTGTGGAGAAATACCTAGTGCATGCTCTGCAAGGAGC	1264
Qy	1264	GTGAGCTCAAGCCAGGCCCCACAGCTGACAGCCTGGCCAGAGCTGGGAGCAGCGGC	1323
Db	1265	GTGAGCTCAAGCCAGGCCCCACAGCTGACAGCCTGGCCAGAGCTGGGAGCAGCGGC	1324
Qy	1324	TCCAGATCCCGGAGCCAGCCCCCAAACTCAGGACAAAGAGGTGCTGCTCACTGAG	1383
Db	1325	TCCAGATCCCGGAGCCAGCCCCCAAACTCAGGACAAAGAGGTGCTGCTCACTGAG	1384
Qy	1384	AAACTCAAGCCAGTGGATTATGATACGAGAAAGTCCACTGGGCCAGCCACAGCTC	1443
Db	1385	AAACTCAAGCCAGTGGATTATGATACGAGAAAGTCCACTGGGCCAGCCACAGCTC	1444
Qy	1444	CGCTTGGCAGAGGCTCTCTCGAGAGGTGCACAGGATGGAGGCAAGCAGCTGGCTTC	1503
Db	1445	CGCTTGGCAGAGGCTCTCTCGAGAGGTGCACAGGATGGAGGCAAGCAGCTGGCTTC	1504
Qy	1504	CAGTGGCTGTCAAAAAGGTGCGCTGGAAGTATTTGGGCGAGAGGCTGATGGCATGT	1563
Db	1505	CAGTGGCTGTCAAAAAGGTGCGCTGGAAGTATTTGGGCGAGAGGCTGATGGCATGT	1564
Qy	1564	GCAGGATTGACTCACCCAGAAATGTCCCTTGTATGGAGCTGTGAGAAAGGCTTGG	1623
Db	1565	GCAGGATTGACTCACCCAGAAATGTCCCTTGTATGGAGCTGTGAGAAAGGCTTGG	1624
Qy	1624	GTCAACATCTTCATGAGCTCTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGAGCAG	1683
Db	1625	GTCAACATCTTCATGAGCTCTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGAGCAG	1684
Qy	1684	GGCTGTCTCCAGAGGACCGGCGCTGTACTACCTGGGCCAGGCGCTGGAGGGTCTGGA	1743

Db	1685	GGCTGTCTCCAGAGGACCGGCGCTGTACTACCTGGGCCAGGCGCTGGAGGGTCTGGA	1744
Qy	1744	TACCTCCACTCACGAAGATTCTCATGGGACGCTCAAAAGCTGACAAAGTGTCTCTGTCC	1803
Db	1745	TACCTCCACTCACGAAGATTCTCATGGGACGCTCAAAAGTGTGACAAAGTGTCTCTGTCC	1804
Qy	1804	AGCATGGAGCCACGAGGCCCTCTGTGACTTTGGCCATGTGTGTCTTTCAACCTGAT	1863
Db	1805	AGCATGGAGCCACGAGGCCCTCTGTGACTTTGGCCATGTGTGTCTTTCAACCTGAT	1864
Qy	1864	GGCTTGGGAAAGTCTTTGCTCACAGGGACTACATCCCTTGGCACAGAGACCAATGCT	1923
Db	1865	GGCTTGGGAAAGTCTTTGCTCACAGGGACTACATCCCTTGGCACAGAGACCAATGCT	1924
Qy	1924	CCGAGGTGGTGTGGGCGAGGCTGCGAGCCCAAGGTGGATGTCTGGAGCAGCTGCTGT	1983
Db	1925	CCGAGGTGGTGTGGGCGAGGCTGCGAGCCCAAGGTGGATGTCTGGAGCAGCTGCTGT	1984
Qy	1984	ATGATGTGCACATGTCTCAAGCGCTGCCACCCCTGGACTCAAGTTCTTCGAGGGCGCTC	2043
Db	1985	ATGATGTGCACATGTCTCAAGCGCTGCCACCCCTGGACTCAAGTTCTTCGAGGGCGCTC	2044
Qy	2044	TGCCTCAAGATTGCCAGGAGCTTCCGCTGTGGAGGAGATCCCACTCTCTGGCCCCCT	2103
Db	2045	TGCCTCAAGATTGCCAGGAGCTTCCGCTGTGGAGGAGATCCCACTCTCTGGCCCCCT	2104
Qy	2104	CTCACGCCCAGGGCATCTCAAGAGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA	2163
Db	2105	CTCACGCCCAGGGCATCTCAAGAGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA	2164
Qy	2164	GCGAGCTGGAGGAAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAAGAGCCCT	2223
Db	2165	GCGAGCTGGAGGAAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAAGAGCCCT	2224
Qy	2224	TGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAATCAAGCCAAATACCAACAG	2283
Db	2225	TGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAATCAAGCCAAATACCAACAG	2284
Qy	2284	ACCTTCATGCCAGCGAGAGCTTTCGCAAGGGCCCCAGGGCCCCCGCCAGCTGAG	2343
Db	2285	ACCTTCATGCCAGCGAGAGCTTTCGCAAGGGCCCCAGGGCCCCCGCCAGCTGAG	2344
Qy	2344	GAGCAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCCAGAGCCA	2403
Db	2345	GAGCAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCCAGAGCCA	2404
Qy	2404	AACAAGTCTCTCTCTTGAATTTTGAGCAAGGAGGTCTGGGATGTGGGAAACCTTACCT	2463
Db	2405	AACAAGTCTCTCTCTTGAATTTTGAGCAAGGAGGTCTGGGATGTGGGAAACCTTACCT	2464
Qy	2464	CTGTCTCTCTGGAGCCAGCCCTGCGCAAGAAACCCAGCTCACAGAGCGGAAAGCAACC	2523
Db	2465	CTGTCTCTCTGGAGCCAGCCCTGCGCAAGAAACCCAGCTCACAGAGCGGAAAGCAACC	2524
Qy	2524	GTCCCGGAGCAGGAACTGCGAGCAGCTGGAATAGAAATTAATTCCTCAACAGAGCTGTCCAG	2583
Db	2525	GTCCCGGAGCAGGAACTGCGAGCAGCTGGAATAGAAATTAATTCCTCAACAGAGCTGTCCAG	2584
Qy	2584	CCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTGCTCTCAGCATGCAAGCTCTCC	2643
Db	2585	CCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTGCTCTCAGCATGCAAGCTCTCC	2644
Qy	2644	CTGTCCGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCCCTG	2703
Db	2645	CTGTCCGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCCCTG	2704
Qy	2704	AGCTCAGCGCTACACTCTCTGGAGCAGCAGGCCCGAGGCTCGAAGCTCCAGCTGGAACATG	2763
Db	2705	AGCTCAGCGCTACACTCTCTGGAGCAGCAGGCCCGAGGCTCGAAGCTCCAGCTGGAACATG	2764
Qy	2764	GTGTGCGCCCGGGCGGCCACCGCACCCCAAGCTATTTCAATGTGTGAAAGTCCAA	2823

PD 02-MAY-2002.
XX PF 15-OCT-2001; 2001EP-00124604.
XX PR 16-OCT-2000; 2000US-0240750P.
XX PA (AXXI-) AXIXIMA PHARM AG.
XX PS Schubart D, Habenberger P, Stein-Gerlach M, Bevec D;
XX PI P-PSDB; AAU80371.
XX DR WPI; 2002-373930/41.
XX DR P-PSDB; AAU80371.
XX PT Identifying agents for treatment or prevention of cytomegalovirus
XX PT infection, comprises contacting test compound with cellular kinase and
XX PT detecting change in cellular kinase activity.
XX PS Disclosure; Page 29-34; 49pp; English.
XX CC The present invention relates to a new method for identifying compounds
XX CC for treating and/or preventing cytomegalovirus (CMV) infection and/or
XX CC related diseases. The method of the invention comprises contacting a test
XX CC compound with at least one of the cellular kinases RICK, RIP, Nck,
XX CC Interacting kinase, MKK3 and SRPK-2 and detecting any change in kinase
XX CC activity. The method of the invention can be used to treat and/or prevent
XX CC CMV infections and related diseases. Oligonucleotides that can detect the
XX CC specified kinases can also be used for diagnosis of infection. The
XX CC present nucleic acid sequence encodes the human cellular kinase NIK (Nck-
XX CC Interacting kinase) protein of the invention, as described above
XX CC
XX SQ Sequence 4596 BP; 1055 A; 1401 C; 1299 G; 841 T; 0 U; 0 Other;
Query Match 99.6%; Score 4578.4; DB 6; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 5 GGGGAGCTGTCCGCTGTGGAACGCTGTAGCTGTGA-AGTGGACTCTGTACCAATTGAGG 63
Db 5 GGGGAGCTGTCCGCTGTGGAACGCTGTAGCTGTGAACGCTGTGTTACCAATTGAGG 64
Qy 64 ATGTTTGGAGATGAGTATGTGTGGCAGAGGACACATAAACAGCGAGACCTTTGGCC 123
Db 65 ATGTTTGGAGATGAGTATGTGTGGCAGAGGACACATAAACAGCGAGACCTTTGGCC 124
Qy 124 CCTGCTTCTCCCAACCAAGCTGACCTGTGTTCTCCAGGTCTGGGATCTAAGT 183
Db 125 CCTGCTTCTCCCAACCAAGCTGACCTGTGTTCTCCAGGTCTGGGATCTAAGT 184
Qy 184 GACCTGCTCTGTGTTGTCTCTCAGGATGAGCACAAGCCTGGGAGTGGCAGTGATG 243
Db 185 GACCTGCTCTGTGTTGTCTCTCAGGATGAGCACAAGCCTGGGAGTGGCAGTGATG 244
Qy 244 GAAATGGCTCCCAAGTGGCTTGTGCTCAGCTGAGCACAAGCCTGGGAGTGGCAGTGATG 303
Db 245 GAAATGGCTCCCAAGTGGCTTGTGCTCAGCTGAGCACAAGCCTGGGAGTGGCAGTGATG 304
Qy 304 CCAAGGAGAGAGACGCGCACTGGGAGAGAAACAGAGCTCCGCTCAAGCTTTGAGGCC 363
Db 305 CCAAGGAGAGAGACGCGCACTGGGAGAGAAACAGAGCTCCGCTCAAGCTTTGAGGCC 364
Qy 364 GTGAGAGAGAGCTTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGACGTGATTACCAAG 423
Db 365 GTGAGAGAGAGCTTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGACGTGATTACCAAG 424
Qy 424 GGCAAGAGAGAGCTTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGACGTGATTACCAAG 483
Db 425 GGCAAGAGAGAGCTTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGACGTGATTACCAAG 484
Qy 484 GAGTGTGAGATACCAAGAGTTCAGGCCACCTTTTCAGAACCGCATTTTCATCGCTGGG 543
Db 485 GAGTGTGAGATACCAAGAGTTCAGGCCACCTTTTCAGAACCGCATTTTCATCGCTGGG 544
Qy 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCAAACAAATGTGGCCCATGCT 603

Db 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCAAACAAATGTGGCCCATGCT 604
Qy 604 ACAGAGGGCAAAATGGCCCGTGTGTTTGAAGGAAAGCGTTCGACGAAAGCCGGGAG 663
Db 605 ACAGAGGGCAAAATGGCCCGTGTGTTTGAAGGAAAGCGTTCGACGAAAGCCGGGAG 664
Qy 664 AAACGGAAGAAGAGAGCTCAAAAGTCCCTGGCTCATGACGAGTGGCTTGGCCAAACCC 723
Db 665 AAACGGAAGAAGAGAGCTCAAAAGTCCCTGGCTCATGACGAGTGGCTTGGCCAAACCC 724
Qy 724 CTCCCGAGGACCTGAGCAGGAGAGTGCACCATCCAGTGCAGGAGGATGAGTCTCA 783
Db 725 CTCCCGAGGACCTGAGCAGGAGAGTGCACCATCCAGTGCAGGAGGATGAGTCTCA 784
Qy 784 CTCCCGGCCCCATATGTTAGAAACACCCCGAGTTCACCAAGCTCTGAAAGAAACACGCG 843
Db 785 CTCCCGGCCCCATATGTTAGAAACACCCCGAGTTCACCAAGCTCTGAAAGAAACACGCG 844
Qy 844 CTTCGGCAACTCTGTTTAAAGCAGCTTGGGAGGCGCTACGCGCGGCTCTGCTCGATCA 903
Db 845 CTTCGGCAACTCTGTTTAAAGCAGCTTGGGAGGCGCTACGCGCGGCTCTGCTCGATCA 904
Qy 904 GAACTCCAAACCTGATCAGCCCTTGCATGTCTGAAACACACGCTGTGAAACTGCAACAC 963
Db 905 GAACTCCAAACCTGATCAGCCCTTGCATGTCTGAAACACACGCTGTGAAACTGCAACAC 964
Qy 964 CCCGAGGAGGAGGCGCCCTGCGCCCGCAGCAGCCCTTCCCTATAGCAGACTGCT 1023
Db 965 CCCGAGGAGGAGGCGCCCTGCGCCCGCAGCAGCCCTTCCCTATAGCAGACTGCT 1024
Qy 1024 CATCCCTTCCCAATCCACCTCTCCAGCCCTGGAACCTCACCTCTGGAGTCTCTCTG 1083
Db 1025 CATCCCTTCCCAATCCACCTCTCCAGCCCTGGAACCTCACCTCTGGAGTCTCTCTG 1084
Qy 1084 GGCATACTGCGCTGTGTAGACAGCCAGAAACCTTGTGCTGACCCACACCTGAGCAAACTG 1143
Db 1085 GGCATACTGCGCTGTGTAGACAGCCAGAAACCTTGTGCTGACCCACACCTGAGCAAACTG 1144
Qy 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCGCGCCACACCTGGAGCCAGCTGCTGTCT 1203
Db 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCGCGCCACACCTGGAGCCAGCTGCTGTCT 1204
Qy 1204 CGTGTGCTCCCATGAGAAAGTCTGTGGAGGAATACCTAGTGCATGTCTGTCAAGGAGC 1263
Db 1205 CGTGTGCTCCCATGAGAAAGTCTGTGGAGGAATACCTAGTGCATGTCTGTCAAGGAGC 1264
Qy 1264 GTGAGCTCAAGCCAGGCGCCACAGCTGACCTGAGCTGGCCAAAGACCTGGGAGCAGCGGGC 1323
Db 1265 GTGAGCTCAAGCCAGGCGCCACAGCTGACCTGAGCCAAAGACCTGGGAGCAGCGGGC 1324
Qy 1324 TCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAACGAGGCTGCTGCTCACTGAG 1383
Db 1325 TCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAACGAGGCTGCTGCTCACTGAG 1384
Qy 1384 AAATCAAGCAGTGGATATGATGATCCGAGAGAAAGTCCACTGGGCGACGCAAGCTC 1443
Db 1385 AAATCAAGCAGTGGATATGATGATCCGAGAGAAAGTCCACTGGGCGACGCAAGCTC 1444
Qy 1444 CGCTTGGGAGAGGCTCTTCGGAGAGGTGCACAGGATGGAGGCAAGAGGAGCTGGCTTC 1503
Db 1445 CGCTTGGGAGAGGCTCTTCGGAGAGGTGCACAGGATGGAGGCAAGAGGAGCTGGCTTC 1504
Qy 1504 CAGTGGCTGTCAAAAGGTGGCGCTGGAAGTATTTTCGGGAGAGGAGCTGATGGCATGT 1563
Db 1505 CAGTGGCTGTCAAAAGGTGGCGCTGGAAGTATTTTCGGGAGAGGAGCTGATGGCATGT 1564
Qy 1564 GCAGGATTTGACCTCACCCAGAAATTTGCTTGTGTATGAGAGTGTGAGAGAGGCGCTTGG 1623
Db 1565 GCAGGATTTGACCTCACCCAGAAATTTGCTTGTGTATGAGAGTGTGAGAGAGGCGCTTGG 1624
Qy 1624 GTCAACATCTTATGAGAGTGTGGAGGTGGCTTCCCTGGGCGCAGCTGGTCAAGGAGCAG 1683

Db 1625 GTCAACATCTTTCATGGAGTGCTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGAGCAG 1684
Qy 1684 GGCTGTCTCCAGAGGACCGGSCCTGTACTACTCTGGGCGCAGGCGCTCGAGGGTCTGCAA 1743
Db 1685 GGCTGTCTCCAGAGGACCGGSCCTGTACTACTCTGGGCGCAGGCGCTCGAGGGTCTGAAA 1744
Qy 1744 TACCTCCATCTCAAGAGGATCTGCATGGGACGTCGTAAGCTGACAAACGTCCTCTGTCC 1803
Db 1745 TACCTCCATCTCAAGAGGATCTGCATGGGACGTCGTAAGCTGACAAACGTCCTCTGTCC 1804
Qy 1804 AGCGATGGGAGGACGCGACCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGAT 1863
Db 1805 AGCGATGGGAGGACGCGACCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGAT 1864
Qy 1864 GSCCTGGGAAGTCTTGTCTACAGGGACTACATCTCCCTGGGCACAGAGACCCACATGGCT 1923
Db 1865 GSCCTGGGAAGTCTTGTCTACAGGGAGTACATCTCCCTGGGCACAGAGACCCACATGGCT 1924
Qy 1924 CCGGAGGTGGTCTGGGCGAGGCTGCGACGCGCAAGGTGGATGCTCTGGAGCAGCTGCTGT 1983
Db 1925 CCGGAGGTGGTCTGGGCGAGGCTGCGACGCGCAAGGTGGATGCTCTGGAGCAGCTGCTGT 1984
Qy 1984 ATGATGCTGCACATGCTCAAGGGCTGCGACCCCTGCACTCAGTCTCTTCGGAGGGCGCTC 2043
Db 1985 ATGATGCTGCACATGCTCAAGGGCTGCGACCCCTGCACTCAGTCTCTTCGGAGGGCGCTC 2044
Qy 2044 TSCCTCAAGATTGCGAGGAGCCTCCGCTGTGAGGGAGATCCACCCCTCTCTGGCCCT 2103
Db 2045 TGCCCTCAAGATTGCGAGGAGCCTCCGCTGTGAGGGAGATCCACCCCTCTCTGGCCCT 2104
Qy 2104 CTCACAGCCCGAGGCCATCAAGAGGGGCTGAGGAAAGGCCCATCCACCGCTGTCTGCA 2163
Db 2105 CTCACAGCCCGAGGCCATCAAGAGGGGCTGAGGAAAGGCCCATCCAACGCTGTCTGCA 2164
Qy 2164 GCGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT 2223
Db 2165 GCGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT 2224
Qy 2224 TGGAGGGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACACACAG 2283
Db 2225 TGGAGGGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACACACAG 2284
Qy 2284 ACCCTCCATGCCAGCGCAGAGAGCTTTGCGCAAGGGGCCCAAGGGCCCGGCGCAGCTGAG 2343
Db 2285 ACCCTCCATGCCAGCGCAGAGAGCTTTGCGCAAGGGGCCCAAGGGCCCGGCGCAGCTGAG 2344
Qy 2344 GAGACAAAGGAGAGCCCTTAAGCTCAGCTCTCTCCACAGAGGCCCCCAGAGGCA 2403
Db 2345 GAGACAAAGGAGAGCCCTTAAGCTCAGCTCTCTCCACAGAGGCCCCCAGAGGCA 2404
Qy 2404 AACAGTCTCTCCCTTGACTTTGAGCAGGAGGAGTCTGGGATGTGGGAACCTTACT 2463
Db 2405 AACAGTCTCTCCCTTGACTTTGAGCAGGAGGAGTCTGGGATGTGGGAACCTTACT 2464
Qy 2464 CTGTCTCTCGAGCAGCCCTGCGAAGACCCAGCTCACAGAGCGGAAAGCAAC 2523
Db 2465 CTGTCTCTCGAGCAGCCCTGCGAAGACCCAGCTCACAGAGCGGAAAGCAAC 2524
Qy 2524 GTCCGGAGCAGAACTGCAGAGCTGGAAATAGAAATTAATCTCAACAGCTGTCCAG 2583
Db 2525 GTCCGGAGCAGAACTGCAGAGCTGGAAATAGAAATTAATCTCAACAGCTGTCCAG 2584
Qy 2584 CCATTTTCTCTGGAGCAGAGGCAATTTCTCTGTGCTCTAGCATCGACAGCTCTCC 2643
Db 2585 CCATTTTCTCTGGAGCAGAGGCAATTTCTCTGTGCTCTAGCATCGACAGCTCTCC 2644
Qy 2644 CTGTCTGATGACAGTGAAGAACCCATCAAGGGCTCTCAAGAGCTCGGGGACACCTG 2703
Db 2645 CTGTCTGATGACAGTGAAGAACCCATCAAGGGCTCTCAAGAGCTCGGGGACACCTG 2704
Qy 2704 AGCTCAGGCGTACATCTCTGGAGCAGCCAGGCGGAGCTCGAAGCTCGAGCTGGAACATG 2763
Db 2705 AGCTCAGGCGTACATCTCTGGAGCAGCCAGGCGGAGCTCGAAGCTCGAAGCTCGAAGCTGGAACATG 2764

Qy 2764 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTGTGAAAGTCCAA 2823
Db 2765 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTGTGAAAGTCCAA 2824
Qy 2824 ATACAGTCTCTTAATGTGGAACACTGTCGAGATCCGGGAGTTCACCGGGTCAAGTGGGA 2883
Db 2825 ATACAGTCTCTTAATGTGGAACACTGTCGAGATCCGGGAGTTCACCGGGTCAAGTGGGA 2884
Qy 2884 GACATCGCACCTGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCAACAA 2943
Db 2885 GACATCGCACCTGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCAACAA 2944
Qy 2944 GACGGCAGCCTGTTGCTACGACATGAGAGTGCAGACTCGGGCATCGACTGCAGTGC 3003
Db 2945 GACGGCAGCCTGTTGCTACGACATGAGAGTGCAGACTCGGGCATCGACTGCAGTGC 3004
Qy 3004 ACATGGCCCTGATGGCAGCTTGCCTGGAGCTGGAGGTCAGACATGGCCAGCTGAG 3063
Db 3005 ACATGGCCCTGATGGCAGCTTGCCTGGAGCTGGAGGTCAGACATGGCCAGCTGAG 3064
Qy 3064 AACAGGCCCTAACCTGCCCTCCACCGCGGCTCCACACTCGCGAAAGCAGCTTCCTG 3123
Db 3065 AACAGGCCCTAACCTGCCCTCCACCGCGGCTCCACACTCGCGAAAGCAGCTTCCTG 3124
Qy 3124 CTCGCTGACGATGCTGCCCTGAAAAACACAGGCTCAGCCGTTCCCGAGGGGATTCAGCC 3183
Db 3125 CTCGCTGACGATGCTGCCCTGAAAAACACAGGCTCAGCCGTTCCCGAGGGGATTCAGCC 3184
Qy 3184 CCCGGCTCAAGTGGGAAACAGGGCTCAGAGCAAGTGGGGGCAAGCAAGTGC 3243
Db 3185 CCCGGCTCAAGTGGGAAACAGGGCTCAGAGCAAGTGGGGGCAAGCAAGTGC 3244
Qy 3244 TCCAGGATTTACACTGAGCCCTGCCACCTGCTGAAACAAACATCCGCCACGTAA 3303
Db 3245 TCCAGGATTTACACTGAGCCCTGCCACCTGCTGAAACAAACATCCGCCACGTAA 3304
Qy 3304 GAGACAGAGGAGATGCGAGAGTTACTGGGAAAAACAAACAGGGATCTTTTCTGCC 3363
Db 3305 GAGACAGAGGAGATGCGAGAGTTACTGGGAAAAACAAACAGGGATCTTTTCTGCC 3364
Qy 3364 CTGCTCAGTTCGAGTTGGCTGACCCCTTGGATCAGTGACCATTTGTTGGCAGACAGG 3423
Db 3365 CTGCTCAGTTCGAGTTGGCTGACCCCTTGGATCAGTGACCATTTGTTGGCAGACAGG 3424
Qy 3424 GGAGCAGCTTCAGCCTGGGTGAGAGGGTGGCGAGCCCTTCGGCCCTCACCCTC 3483
Db 3425 GGAGCAGCTTCAGCCTGGGTGAGAGGGTGGCGAGCCCTTCGGCCCTCACCCTC 3484
Qy 3484 CAGGCTGCTGAGAGTGTCAAGTGTGTAAGGGGCCCAAACCTCAGTTTCAGTGCAGAACCA 3543
Db 3485 CAGGCTGCTGAGAGTGTCAAGTGTGTAAGGGGCCCAAACCTCAGTTTCAGTGCAGAACCA 3544
Qy 3544 GGTTCAGAGGTATGCCCGCCGTTAAGTTAAGGGGCCCTCTAAACCCCTTGCCTGGCCTC 3603
Db 3545 GGTTCAGAGGTATGCCCGCCGTTAAGTTAAGGGGCCCTCTAAACCCCTTGCCTGGCCTC 3604
Qy 3604 ACCTGGCCAGCTCACCCCTTTGGGTGAGGGGAAAGAAATGCGCTGAGCCCTGGGAGGCT 3663
Db 3605 ACCTGGCCAGCTCACCCCTTTGGGTGAGGGGAAAGAAATGCGCTGAGCCCTGGGAGGCT 3664
Qy 3664 CCCTGGTGAATATACACACATCTTTTCAGTTGTTGCAACACAGGTCCTGAGTTGACCTCT 3723
Db 3665 CCCTGGTGAATATACACACATCTTTTCAGTTGTTGCAACACAGGTCCTGAGTTGACCTCT 3724
Qy 3724 GGTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGGTTCAGTCCCGCAGCATGT 3783
Db 3725 GGTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGGTTCAGTCCCGCAGCATGT 3784
Qy 3784 GCCCTTTGCTGCTGCTACCATCTTTCCAGAGCAGCAGGCCCCCGAGCCCTTCAGG 3843
Db 3785 GCCCTTTGCTGCTGCTACCATCTTTCCAGAGCAGCAGGCCCCCGAGCCCTTCAGG 3844

Dd 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCAACAAATGTGGCCCATGCT 604
Qy 604 ACAGAGGCAAAATGSCCGTGTGTGTGGAGAGGAAAGCGTCGAGCAAAAGCCCCGGAAG 663
Dd 605 ACAGAGGCAAAATGSCCGTGTGTGTGGAGAGGAAAGCGTCGAGCAAAAGCCCCGGAAG 664
Qy 664 AAACGGAGAGAAAGAGCTCAAAAGTCCCTGGCTCATGCAAGAGTGGCCTTGCCCAAAAGCC 723
Dd 665 AAACGGAGAGAAAGAGCTCAAAAGTCCCTGGCTCATGCAAGAGTGGCCTTGCCCAAAAGCC 724
Qy 724 CTCCCAAGAGCCCTTGAGCAGAGAGTGCACCAATCCCAAGTGCAGAGAGATGAGTCTCA 783
Dd 725 CTCCCAAGAGCCCTTGAGCAGAGAGTGCACCAATCCCAAGTGCAGAGAGATGAGTCTCA 784
Qy 784 CTCGGGGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCCTCTGAAGAAACAGGC 843
Dd 785 CTCGGGGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCCTCTGAAGAAACAGGC 844
Qy 844 CTTGGGCAACTCTGTTTTAAGCAGCTTGGCGAGGGCCTACGGCCGGCTCTGCCTCGATCA 903
Dd 845 CTTGGGCAACTCTGTTTTAAGCAGCTTGGCGAGGGCCTACGGCCGGCTCTGCCTCGATCA 904
Qy 904 GAACTCCAAACTGATCAGCCCTTGCAATGCTGAAACCAAGTGTGAAACTGCACCAAC 963
Dd 905 GAACTCCAAACTGATCAGCCCTTGCAATGCTGAAACCAAGTGTGAAACTGCACCAAC 964
Qy 964 CCCCAGGAGGAGGCCCTGCCCTGCCACGACCCCTTCCCTATAGCAGATGCTCT 1023
Dd 965 CCCCAGGAGGAGGCCCTGCCCTGCCACGACCCCTTCCCTATAGCAGATGCTCT 1024
Qy 1024 CATCCCTTCCCAATCCACCTCTCCAGCCCTGGAAACCTCACCCCTCTGGAGTCCCTTCTG 1083
Dd 1025 CATCCCTTCCCAATCCACCTCTCCAGCCCTGGAAACCTCACCCCTCTGGAGTCCCTTCTG 1084
Qy 1084 GGCAAACTGGCTGTGTAGACAGCAGAAACCTTGCCTGACCCCACTGTAGCAAACTG 1143
Dd 1085 GGCAAACTGGCTGTGTAGACAGCAGAAACCTTGCCTGACCCCACTGTAGCAAACTG 1144
Qy 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTTGGAGCCAGCTGCTGTCT 1203
Dd 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTTGGAGCCAGCTGCTGTCT 1204
Qy 1204 CGTGTGTCCCATGAGAAGTTTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGCAGC 1263
Dd 1205 CGTGTGTCCCATGAGAAGTTTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGCAGC 1264
Qy 1264 GTGAGTCAAGCCAGGCCACAGCCTGACAGCCTGGCCAAAGCACTGTGGCAGCAAGGGGC 1323
Dd 1265 GTGAGTCAAGCCAGGCCACAGCCTGACAGCCTGGCCAAAGCACTGTGGCAGCAAGGGGC 1324
Qy 1324 TCCAGATCCCGGAGCCAGCCCAAACTGAGGACACAGGGGTGCTGCTCACTGAG 1383
Dd 1325 TCCAGATCCCGGAGCCAGCCCAAACTGAGGACACAGGGGTGCTGCTCACTGAG 1384
Qy 1384 AAACCTCAAGCCAGTGGATTATGAGTACCGAGAAGAAAGTCCACTGGGCCACGACCAAGCTC 1443
Dd 1385 AAACCTCAAGCCAGTGGATTATGAGTACCGAGAAGAAAGTCCACTGGGCCACGACCAAGCTC 1444
Qy 1444 CGCCTGGCAGAGGCTCTCTCGAGAGGTGCAAGAGTGGAGCAAGCAGACTGGCTTC 1503
Dd 1445 CGCCTGGCAGAGGCTCTCTCGAGAGGTGCAAGAGTGGAGCAAGCAGACTGGCTTC 1504
Qy 1504 CAGTGGCTCAAAAAGGTGGCCTGGAAGTATTTCCGGCAGAGGAGCTGATGCAATGT 1563
Dd 1505 CAGTGGCTGTCAAAAAGGTGGCCTGGAAGTATTTCCGGCAGAGGAGCTGATGCAATGT 1564
Qy 1564 GCAGGATTGACCTCAACCAAGATTGTCCCTTTGTATGAGTGTGTGAGAGAAGGGCTTGG 1623
Dd 1565 GCAGGATTGACCTCAACCAAGATTGTCCCTTTGTATGAGTGTGTGAGAGAAGGGCTTGG 1624
Qy 1624 GTCAACATCTCATGAGGTGTGGAAGGTGGCTCCCTGGGCCAGCTGTGTCAAGAGCAG 1683
Dd 1625 GTCAACATCTCATGAGGTGTGGAAGGTGGCTCCCTGGGCCAGCTGTGTCAAGAGCAG 1684

Qy 1684 GGCTGTCTCCAGAGGACCGGGCCCTGTGTACTCTGGGCCAGGGCCCTGGAGGGTCTGGAA 1743
Dd 1685 GGCTGTCTCCAGAGGACCGGGCCCTGTGTACTCTGGGCCAGGGCCCTGGAGGGTCTGGAA 1744
Qy 1744 TACCTCCACTCAACGAAGGATTTCTGCATGGGACGFTCAAAGCTGACAAAGTGTCTCTGTCC 1803
Dd 1745 TACCTCCACTCAACGAAGGATTTCTGCATGGGACGFTCAAAGCTGACAAAGTGTCTCTGTCC 1804
Qy 1804 AGCGATGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGTGTCTTCAACCTGAT 1863
Dd 1805 AGCGATGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGTGTCTTCAACCTGAT 1864
Qy 1864 GSCCTGGGAAAGTCTCTCTCACAGGGGACTACATCTCTGGCCACAGAGACCCACATGGCT 1923
Dd 1865 GSCCTGGGAAAGTCTCTCTCACAGGGGACTACATCTCTGGCCACAGAGACCCACATGGCT 1924
Qy 1924 CCGGAGGTGTGTCTGGGCAAGGAGTGCAGACCCCAAGGTGTGTCTGGAGCAGCTGTCT 1983
Dd 1925 CCGGAGGTGTGTCTGGGCAAGGAGTGCAGACCCCAAGGTGTGTCTGGAGCAGCTGTCT 1984
Qy 1984 ATGATGTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTTCGAGGGCCGCTC 2043
Dd 1985 ATGATGTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTTCGAGGGCCGCTC 2044
Qy 2044 TGCTTCAAGATTGCCAGCGAGCTCCGCTGTGTAGGGAGATCCCAACCTCTCTGCCGCCCT 2103
Dd 2045 TGCTTCAAGATTGCCAGCGAGCTCCGCTGTGTAGGGAGATCCCAACCTCTCTGCCGCCCT 2104
Qy 2104 CTCACAGCCCAGGCCATCCAGAGGGGTGAGGAAAGGCCATCCACCGGTGTCTGCA 2163
Dd 2105 CTCACAGCCCAGGCCATCCAGAGGGGTGAGGAAAGAGCCATCCACCGGTGTCTGCA 2164
Qy 2164 GCGAGCTTGGAGGGAAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAGAGCCCT 2223
Dd 2165 GCGAGCTTGGAGGGAAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAGAGCCCT 2224
Qy 2224 TGGAGGGAGAAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACCACAG 2283
Dd 2225 TGGAGGGAGAAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACCACAG 2284
Qy 2284 ACCCTTCCATGCCACGCGAGAGAGCTTTCCGCCAAGGGCCCAAGGCCCGGCCAGCTGAG 2343
Dd 2285 ACCCTTCCATGCCACGCGAGAGAGCTTTCCGCCAAGGGCCCAAGGGCCCGGCCAGCTGAG 2344
Qy 2344 GAGACAAAGCAGAGAGCCCTTAAGCTCCAGCTCTCTCTCCACAGAGGCCCCCAGAGCA 2403
Dd 2345 GAGACAAAGCAGAGAGCCCTTAAGCTCCAGCTCTCTCTCCACAGAGGCCCCCAGAGCA 2404
Qy 2404 AACAGTCTCTCTCTTCACTTTTGAAGCAAGAGAGTCTGGGATGTGGAAACCTTACCT 2463
Dd 2405 AACAGTCTCTCTCTTCACTTTTGAAGCAAGAGAGTCTGGGATGTGGAAACCTTACCT 2464
Qy 2464 CTGTCTCTCTGGAGCCAGCCCTTGCAGAAAACCCAGCTCACAGAGCGGAAAGCAACC 2523
Dd 2465 CTGTCTCTCTGGAGCCAGCCCTTGCAGAAAACCCAGCTCACAGAGCGGAAAGCAACC 2524
Qy 2524 GTCCCGGAGCAGGAATCTGAGCAGCTGGAATATGAAATTTATTCCTCAAAGCTGTCTCCAG 2583
Dd 2525 GTCCCGGAGCAGGAATCTGAGCAGCTGGAATATGAAATTTATTCCTCAAAGCTGTCTCCAG 2584
Qy 2584 CCAATTTCTCTGGAGGACGAGCAAAATTTCTCTGCTCTCAGCATCGACAGCTCTCTCC 2643
Dd 2585 CCAATTTCTCTGGAGGACGAGCAAAATTTCTCTGCTCTCAGCATCGACAGCTCTCTCC 2644
Qy 2644 CTGTCCGATGACAGTGAAGAAACCATCAAAGGCTCTCAAAGCTTCGCGGAGCAACCTGT 2703
Dd 2645 CTGTCCGATGACAGTGAAGAAACCATCAAAGGCTCTCAAAGCTTCGCGGAGCAACCTGT 2704
Qy 2704 AGCTCAGGGTACTCTCTGGAGCAGCCAGGCTCGAAGCTTCAGCTCGGAGCAATG 2763
Dd 2705 AGCTCAGGGTACTCTCTGGAGCAGCCAGGCTCGAAGCTTCAGCTCGGAGCAATG 2764

```
QY 2764 GTGCTGGCCGGGGGGGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAGTCCAA 2823
DB 2765 GTGCTGGCCGGGGGGGGGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAGTCCAA 2824
QY 2824 ATACAGTCTCTTAATGGTGAACACACTGACATCCGGGAGTTCCACCGGGTCAAGTGGGA 2883
DB 2825 ATACAGTCTCTTAATGGTGAACACACTGACATCCGGGAGTTCCACCGGGTCAAGTGGGA 2884
QY 2884 GACATCGGCACCTGGCATCAGCAGCCAGATCCAGCTGCGAGCTTCCAGCTTGGTCAACCAA 2943
DB 2885 GACATCGGCACCTGGCATCAGCAGCCAGATCCAGCTGCGAGCTTCCAGCTTGGTCAACCAA 2944
QY 2944 GACGGGACGCTGTTGCTACGACATGAGAGTGGCCAGACTGGGATCGACCTGCGAGTGC 3003
DB 2945 GACGGGACGCTGTTGCTACGACATGAGAGTGGCCAGACTGGGATCGACCTGCGAGTGC 3004
QY 3004 ACACCTGGCCCTTGATGGGAGCTTCCGCTGGAGCTGAGAGGTCAAGCATGGCCAGCTGGAG 3063
DB 3005 ACACCTGGCCCTTGATGGGAGCTTCCGCTGGAGCTGAGAGGTCAAGCATGGCCAGCTGGAG 3064
QY 3064 AACAGGCCCTAAACCTGCCCCCTCACCGCCGGCTCCACACTGCGGAAAGCAGCCTTCCCTG 3123
DB 3065 AACAGGCCCTAAACCTGCCCCCTCACCGCCGGCTCCACACTGCGGAAAGCAGCCTTCCCTG 3124
QY 3124 CTCGCTGACAGATGCTGCCCCCTGAAACACAGGCTCAGCCGTTCCAGGGGATTTGCCAGCC 3183
DB 3125 CTCGCTGACAGATGCTGCCCCCTGAAACACAGGCTCAGCCGTTCCAGGGGATTTGCCAGCC 3184
QY 3184 CCCCGGCTCACAGTGGGAAACAGGCGCTCGCAGCAGCAAGTGGGGCAAGCAGATGCC 3243
DB 3185 CCCCGGCTCACAGTGGGAAACAGGCGCTCGCAGCAGCAAGTGGGGCAAGCAGATGCC 3244
QY 3244 TCCAGGATTTTACACCTGAGCCCTGCCCCCAGCCCTGCTGAAACCAATCCCGCAGCTGAA 3303
DB 3245 TCCAGGATTTTACACCTGAGCCCTGCCCCCAGCCCTGCTGAAACCAATCCCGCAGCTGAA 3304
QY 3304 GAGACAGAGGAGATGCGAGAGTTACTTGGGAAACAAACAGGAGATCTTTTCTGCC 3363
DB 3305 GAGACAGAGGAGATGCGAGAGTTACTTGGGAAACAAACAGGAGATCTTTTCTGCC 3364
QY 3364 CCTGCTCAGTGCAGTTCGCTGACCCGCTTGATCAGTGACCAATTTGTTGGCAGACAGG 3423
DB 3365 CCTGCTCAGTGCAGTTCGCTGACCCGCTTGATCAGTGACCAATTTGTTGGCAGACAGG 3424
QY 3424 GGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGCGAGCCCTTCGCGCCCTCACCCCTC 3483
DB 3425 GGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGCGAGCCCTTCGCGCCCTCACCCCTC 3484
QY 3484 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGGCCAAACTCAGGTTCAAGTGCAGAACCA 3543
DB 3485 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGGCCAAACTCAGGTTCAAGTGCAGAACCA 3544
QY 3544 GGTACAGAGGTATGCCCCCGGTAGGTTAAGGGGGCCCTCTAAACCCCTTCGCTGGCCCTC 3603
DB 3545 GGTACAGAGGTATGCCCCCGGTAGGTTAAGGGGGCCCTCTAAACCCCTTCGCTGGCCCTC 3604
QY 3604 ACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAGAAATGCCCTGAGCCCTGGGAAGCT 3663
DB 3605 ACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAGAAATGCCCTGAGCCCTGGGAAGCT 3664
QY 3664 CCCTGGTGAATACACCACTTTTTCAGGTTGTTGCAACACAGAGTCTCTGAGTTCACCTCT 3723
DB 3665 CCCTGGTGAATACACCACTTTTTCAGGTTGTTGCAACACAGAGTCTCTGAGTTCACCTCT 3724
QY 3724 GGTTCAGCCAGGACCAAGAGGTGTGTAAGTGAAGTGTCTCAGTCCCGCAGCATGT 3783
DB 3725 GGTTCAGCCAGGACCAAGAGGTGTGTAAGTGAAGTGTCTCAGTCCCGCAGCATGT 3784
QY 3784 GCCCTTGTGCTGGCTTACCACTCTTCCAGAGCAGCAGGCCCGGAGCCCTTTCAGGC 3843
DB 3785 GCCCTTGTGCTGGCTTACCACTCTTCCAGAGCAGCAGGCCCGGAGCCCTTTCAGGC 3844
QY 3844 CCAGCACTGCCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTGTAT 3903
```

RESULT 5

ADR25114

ID ADR25114 standard; DNA; 4596 BP.

XX ADR25114;

AC ADR25114;

DT 21-OCT-2004 (first entry)

XX

DE Breast cancer prognosis marker #975.

XX

KW ds; breast cancer; prognosis; gene expression; diagnosis.

XX

OS Homo sapiens.

XX

XX WO2004065545-A2.

XX

XX 05-AUG-2004.

XX

XX 15-JAN-2004; 2004WO-US001100.

XX

XX 15-JAN-2003; 2003US-00342887.

XX

```
DB 3845 CCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTGTAT 3904
QY 3904 GCAGGATATGCTTCACAGGAAACAGTCTGTGATGGACATGATCAGTGTAAAGGAAAGCAG 3963
DB 3905 GCAGGATATGCTTCACAGGAAACAGTCTGTGATGGACATGATCAGTGTAAAGGAAAGCAG 3964
QY 3964 CAGAGAGAGAGCTCCGGCGCCCCAGCCCACTATCAGTGTCCAGCGTCTGGTTCCCCAG 4023
DB 3965 CAGAGAGAGAGCTCCGGCGCCCCAGCCCACTATCAGTGTCCAGCGTCTGGTTCCCCAG 4024
QY 4024 AGCAGAGCTCAGCATCAGACTGACACTCACCTGCCCTGCCCTGGCCAGAGGGTACTGTC 4083
DB 4025 AGCAGAGCTCAGCATCAGACTGACACTCACCTGCCCTGCCCTGGCCAGAGGGTACTGTC 4084
QY 4084 CGAGCGCACTTTGACATCTGATGACCTCAAGAGCACTTTTCATGGCTGCCCTCTGGCAGGGC 4143
DB 4085 CGAGCGCACTTTGACATCTGATGACCTCAAGAGCACTTTTCATGGCTGCCCTCTGGCAGGGC 4144
QY 4144 AGGCGAGGCGAGTGCACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGACACA 4203
DB 4145 AGGCGAGGCGAGTGCACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGACACA 4204
QY 4204 GTCTTGAGCTGTCCACATGCACTGCTCCTCAAACCTTCCAGATTTCTTAAGAATA 4263
DB 4205 GTCTTGAGCTGTCCACATGCACTGCTCCTCAAACCTTCCAGATTTCTTAAGAATA 4264
QY 4264 GCACCCCTTCCCATTTGCCAGCTTAGCCCTTTCCTCCAGGGGAGCTCTCAGGACTC 4323
DB 4265 GCACCCCTTCCCATTTGCCAGCTTAGCCCTTTCCTCCAGGGGAGCTCTCAGGACTC 4324
QY 4324 ACGTAGCATTTAAATCAGCTGTGAATCGTCAGGGGGTGTCTGCTAGCCCTCAACCTCCTGGG 4383
DB 4325 ACGTAGCATTTAAATCAGCTGTGAATCGTCAGGGGGTGTCTGCTAGCCCTCAACCTCCTGGG 4384
QY 4384 GCAGGGAGCGCCGAGACTTCCTGGGAGAGCTCATTTCCACATCTTGGCCAGACAGACCTT 4443
DB 4385 GCAGGGAGCGCCGAGACTTCCTGGGAGAGCTCATTTCCACATCTTGGCCAGACAGACCTT 4444
QY 4444 TGTCCAGCTGTCCACATTTGAGTTCAGCTGTCTCCCGGGGAGAGAGCCCGGCCCCAGCAC 4503
DB 4445 TGTCCAGCTGTCCACATTTGAGTTCAGCTGTCTCCCGGGGAGAGAGCCCGGCCCCAGCAC 4504
QY 4504 ATAAAGAACTGCAGCCCTTGGTACTGTCAGAGTCTGGGTTGTAGAGAACTCTTTGTAAAGCAA 4563
DB 4505 ATAAAGAACTGCAGCCCTTGGTACTGTCAGAGTCTGGGTTGTAGAGAACTCTTTGTAAAGCAA 4564
QY 4564 TAAAGTTTGGGGTGAATGACAAATGTTAAAAA 4595
DB 4565 TAAAGTTTGGGGTGAATGACAAATGTTAAAAA 4596
```

RESULT 5

ADR25114

ID ADR25114 standard; DNA; 4596 BP.

XX ADR25114;

AC ADR25114;

DT 21-OCT-2004 (first entry)

XX

DE Breast cancer prognosis marker #975.

XX

KW ds; breast cancer; prognosis; gene expression; diagnosis.

XX

OS Homo sapiens.

XX

XX WO2004065545-A2.

XX

XX 05-AUG-2004.

XX

XX 15-JAN-2004; 2004WO-US001100.

XX

XX 15-JAN-2003; 2003US-00342887.

XX

XX	(ROSE-) ROSETTA INPHARMATICS LLC.
PA	(NECA-) NETHERLANDS CANCER INST.
PA	
XX	Van't Veer LJ, He Y;
PI	
XX	WPI; 2004-593473/57.
DR	
XX	Classifying a breast cancer patient according to prognosis comprises
PT	determining the similarity between the level of expression of each of
PT	five genes in a cell sample taken from patient, to control levels.
XX	
XX	Disclosure; SEQ ID NO 975; 226pp; English.
PS	
XX	The invention relates to a method of classifying a breast cancer patient
CC	according to prognosis by determining the similarity between the level of
CC	expression of each of five genes for which markers are listed in the
CC	specification, in a cell sample taken from the breast cancer patient, to
CC	control levels of expression for each respective five genes to obtain a
CC	patient similarity value. The methods are useful for classifying a breast
CC	cancer patient according to prognosis. Kits and computer program products
CC	are useful for data analysis using the diagnostic, prognostic and
CC	statistical methods of the invention. This sequence corresponds to a
CC	marker used in the method of the invention.
XX	
SQ	Sequence 4596 BP; 1055 A; 1401 C; 1299 G; 841 T; 0 U; 0 Other;
Query Match 99.6%; Score 4578.4; DB 13; Length 4596;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1	
QY	5 GGGGACTGTGCGGTGGAACTGTAGCTGTGA-AGGTGACTCTGTACCATTTGAGG 63
DB	5 GGGGACTGTGCGGTGGAACTGTAGCTGTGAGAGTGGACTCTGTACCATTTGAGG 64
QY	64 ATGTTTGGAGATGAGTATGTGTGCGAGGACACATAAACAGCAGAGACCTTTGCC 123
DB	65 ATGTTTGGAGATGAGTATGTGTGCGAGGACACATAAACAGCAGAGACCTTTGCC 124
QY	124 CTTGCTTTTCCCCAACCCAGGCTGACCTGTCTTCTCCAGGTCTGGGATTTCTAAGT 183
DB	125 CTTGCTTTTCCCCAACCCAGGCTGACCTGTCTTCTCCAGGTCTGGGATTTCTAAGT 184
QY	184 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAAGCTTGGAGATGCGATGATG 243
DB	185 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAAGCTTGGAGATGCGATGATG 244
QY	244 GAAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGCGACAGAGAACTCCCCAAG 303
DB	245 GAAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGCGACAGAGAACTCCCCAAG 304
QY	304 CCAAGAGGAGAGACGCCCGCACTGGGGAGAAACAGAGCTCCGTCTACAAGCTTTGAGGCC 363
DB	305 CCAAGAGGAGAGACGCCCGCACTGGGGAGAAACAGAGCTCCGTCTACAAGCTTTGAGGCC 364
QY	364 GTGGAGAGAGCCCTGTGTTTTCGGGAAAGTGGAGATCCTGAATGACGTGATTACCAG 423
DB	365 GTGGAGAGAGCCCTGTGTTTTCGGGAAAGTGGAGATCCTGAATGACGTGATTACCAG 424
QY	424 GGCACAGCCCAAGGAGGCTCCGAGGAGGCGCAGCTGCCATCTCTATCATCGCCAGGCT 483
DB	425 GGCACAGCCCAAGGAGGCTCCGAGGAGGCGCAGCTGCCATCTCTATCATCGCCAGGCT 484
QY	484 GAGTGTGAGATAGCCAGAGTTTACGCCCACTTTTCAGAACGCAATTTTCATCGCTGGG 543
DB	485 GAGTGTGAGATAGCCAGAGTTTACGCCCACTTTTCAGAACGCAATTTTCATCGCTGGG 544
QY	544 TCCAAACAGTACAGCCAGTCCGAGGTCTTGATTCAGATCCCAACAAATGTGGCCCAATGCT 603
DB	545 TCCAAACAGTACAGCCAGTCCGAGGTCTTGATTCAGATCCCAACAAATGTGGCCCAATGCT 604
QY	604 ACAGAGGGCAAAATGGCCCGTGTGTGTTTGGAGGGGAAAGCGTCGACGAGCAAGCCGGAAG 663

QY 1744 TACCTCCACTCAGGAAGATTCTGCATGGGAGCGTCAAAGCTGACAAAGTGTCTCTGTCC 1803
DB 1745 TACCTCCACTCAGGAAGATTCTGCATGGGAGCGTCAAAGCTGACAAAGTGTCTCTGTCC 1804
QY 1804 AGCGATGGAGCCACGACGCCCTCTGTGACTTTGGCCCATGTGTGTCTTTCAACCTGAT 1863
DB 1805 AGCGATGGAGCCACGACGCCCTCTGTGACTTTGGCCCATGTGTGTCTTTCAACCTGAT 1864
QY 1864 GSCCTGGGAAAGTCTTGTCTCACAGGGACTACATCCCTGGGCACAGAGACCCACATGGCT 1923
DB 1865 GSCCTGGGAAAGTCTTGTCTCACAGGGACTACATCCCTGGGCACAGAGACCCACATGGCT 1924
QY 1924 CCGAGGTGTGTCTGGGACGAGCTGCGACGCGCAAGGTGGATGTCTGAGCAGCTGTGT 1983
DB 1925 CCGAGGTGTGTCTGGGACGAGCTGCGACGCGCAAGGTGGATGTCTGAGCAGCTGTGT 1984
QY 1984 ATGATGCTGCACATGCTCAACGGCTCCACCCCTGGACTCAGTTCTTCGGAGGCCGCTC 2043
DB 1985 ATGATGCTGCACATGCTCAACGGCTCCACCCCTGGACTCAGTTCTTCGGAGGCCGCTC 2044
QY 2044 TGCTTCAAGATTGCCAGGAGCTCCGCTGTGAGGGAGATCCCAACCTCTCTGGCCCT 2103
DB 2045 TGCTTCAAGATTGCCAGGAGCTCCGCTGTGAGGGAGATCCCAACCTCTCTGGCCCT 2104
QY 2104 CTACAGCCCGAGCCATCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2163
DB 2105 CTACAGCCCGAGCCATCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2164
QY 2164 GCGAGCTGGGAGGAGGTGAAACCGGCACTACAGCAAGTGGAGTCTGAAAGGCCCT 2223
DB 2165 GCGAGCTGGGAGGAGGTGAAACCGGCACTACAGCAAGTGGAGTCTGAAAGGCCCT 2224
QY 2224 TGGAGGGAGAAATAAAGAAACCAAGACATCCACCGCAAAATCAAGCAATTTACCACAG 2283
DB 2225 TGGAGGGAGAAATAAAGAAACCAAGACATCCACCGCAAAATCAAGCAATTTACCACAG 2284
QY 2284 ACCCTTCCATGCCAGCCGAGAGGCTTTGCGCAAGGGCCCGAGGCCCGGCGCAGCTGAG 2343
DB 2285 ACCCTTCCATGCCAGCCGAGAGGCTTTGCGCAAGGGCCCGAGGCCCGGCGCAGCTGAG 2344
QY 2344 GAGCAACAGGAGAGGCCCTTAAGCTCAGGCTCTCTCTCCACAGAGCCCGGCGCAGAGCCA 2403
DB 2345 GAGCAACAGGAGAGGCCCTTAAGCTCAGGCTCTCTCTCCACAGAGCCCGGCGCAGAGCCA 2404
QY 2404 AACAGTCTCTCCCTTGAAGTCTGAGCAAGGAGTCTGGGATGTGGAAACCTTTACCT 2463
DB 2405 AACAGTCTCTCCCTTGAAGTCTGAGCAAGGAGTCTGGGATGTGGAAACCTTTACCT 2464
QY 2464 CTGTCCTCTCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2523
DB 2465 CTGTCCTCTCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2524
QY 2524 GTCCGGAGCAGGAACTCAGCAGCTGGAATAGAAATTAATTCCTCAACAGCCTGTCCAG 2583
DB 2525 GTCCGGAGCAGGAACTCAGCAGCTGGAATAGAAATTAATTCCTCAACAGCCTGTCCAG 2584
QY 2584 CCAATTTCTCTGGAGGACAGGCAATTTCTCTGTCCTCAGCATCGAGCCTCTCC 2643
DB 2585 CCAATTTCTCTGGAGGACAGGCAATTTCTCTGTCCTCAGCATCGAGCCTCTCC 2644
QY 2644 CTGTCGATGACAGTGAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCCCTG 2703
DB 2645 CTGTCGATGACAGTGAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCCCTG 2704
QY 2704 AGCTCAGCGCTACACTCTCTGGAGCAGCCAGGCTCGAAGCTCAGCTGCGGAAATG 2763
DB 2705 AGCTCAGCGCTACACTCTCTGGAGCAGCCAGGCTCGAAGCTCAGCTGCGGAAATG 2764
QY 2764 GTGTCGGCCCGGGGCGGCCACCGACACCCCAAGCTAATTTCAATGTTGTAAGTCCAA 2823
DB 2765 GTGTCGGCCCGGGGCGGCCACCGACACCCCAAGCTAATTTCAATGTTGTTGTAAGTCCAA 2824

QY 2824 ATACAGTCTCTTAATGTGTGAACACCTGCACATCCGGAGTTTCCACCGGGTCAAAGTGGGA 2883
DB 2825 ATACAGTCTCTTAATGTGTGAACACCTGCACATCCGGAGTTTCCACCGGGTCAAAGTGGGA 2884
QY 2884 GACATCCGCATCTGCATCAGCAGCAGATCCAGCTGCAGCTTTCAGCTTGGTCAACCAA 2943
DB 2885 GACATCCGCATCTGCATCAGCAGCAGATCCAGCTGCAGCTTTCAGCTTGGTCAACCAA 2944
QY 2944 GACGGGAGCCTGTTTCGCTACGACATGGAGGTGCAGACTCGGGCATCGACTGCAAGTGC 3003
DB 2945 GACGGGAGCCTGTTTCGCTACGACATGGAGGTGCAGACTCGGGCATCGACTGCAAGTGC 3004
QY 3004 ACATGCGCCCTGATGGAGCTTCCCTGGAGCTGAGGGTCAAGCATGCGCAGCTGAG 3063
DB 3005 ACATGCGCCCTGATGGAGCTTCCCTGGAGCTGAGGGTCAAGCATGCGCAGCTGAG 3064
QY 3064 AACAGGCCCTAACCTGCCCTCCACCGCGCTCCACATCGCGGAAGAGCCTTCTCTG 3123
DB 3065 AACAGGCCCTAACCTGCCCTCCACCGCGCTCCACATCGCGGAAGAGCCTTCTCTG 3124
QY 3124 CTCGCTGCAGATGCTGCCCTGAAACACAGGCTCAGCGCTTCCAGGGGATTTGCAGCC 3183
DB 3125 CTCGCTGCAGATGCTGCCCTGAAACACAGGCTCAGCGCTTCCAGGGGATTTGCAGCC 3184
QY 3184 CCCCGCTCAGTGGAAACAGGGCTTCGAGCAGCAAGTGGGGCAAGCAAGTGC 3243
DB 3185 CCCCGCTCAGTGGAAACAGGGCTTCGAGCAGCAAGTGGGGCAAGCAAGTGC 3244
QY 3244 TCCAGAGATTTCACCTGAGCCCTGCCACCTGCTGAAACACATCCCGCAGCTGAA 3303
DB 3245 TCCAGAGATTTCACCTGAGCCCTGCCACCTGCTGAAACACATCCCGCAGCTGAA 3304
QY 3304 GAGACAGAGGAGATGGCAGGATTAACCTGGGAAACAAACAGGATCTTTTCTGCC 3363
DB 3305 GAGACAGAGGAGATGGCAGGATTAACCTGGGAAACAAACAGGATCTTTTCTGCC 3364
QY 3364 CCTGCTCCAGTTCGAGTGGCTGACCCGCTTGGATCAGTGACCAATTTTGGCAGACAGG 3423
DB 3365 CCTGCTCCAGTTCGAGTGGCTGACCCGCTTGGATCAGTGACCAATTTTGGCAGACAGG 3424
QY 3424 GGAGAGAGCTTCAGCCTGGGTGAGAAAGGGTGGGAGGCCCTTCGGCCCTCACCCTC 3483
DB 3425 GGAGAGAGCTTCAGCCTGGGTGAGAAAGGGTGGGAGGCCCTTCGGCCCTCACCCTC 3484
QY 3484 CAGCTCTGTGAGAGTCAAGTGTAAAGGGCCCAAACTCAGGTTTCAGTGCAGAACCA 3543
DB 3485 CAGCTCTGTGAGAGTCAAGTGTAAAGGGCCCAAACTCAGGTTTCAGTGCAGAACCA 3544
QY 3544 GGTGAGAGGTATGCCCGCCGCTAGGTTAAAGGGGCCCTCTAAACCCCTTCCTGGCCTC 3603
DB 3545 GGTGAGAGGTATGCCCGCCGCTAGGTTAAAGGGGCCCTCTAAACCCCTTCCTGGCCTC 3604
QY 3604 ACCTGGCAGCTCACCCTTTTGGGTGAGGGGAAAGAAATGCTGACCTCGGAAAGCT 3663
DB 3605 ACCTGGCAGCTCACCCTTTTGGGTGAGGGGAAAGAAATGCTGACCTCGGAAAGCT 3664
QY 3664 CCCTGTTAGAAATACACACTTTTTCAGGTTGTTGCAACACAGCTCTGAGTTGACCTCT 3723
DB 3665 CCCTGTTAGAAATACACACTTTTTCAGGTTGTTGCAACACAGCTCTGAGTTGACCTCT 3724
QY 3724 GGTTCAGCAAGAACCAAGAGGTGTAAAGTGGTTCCTCAGTCCCGCAGACATGT 3783
DB 3725 GGTTCAGCAAGAACCAAGAGGTGTAAAGTGGTTCCTCAGTCCCGCAGACATGT 3784
QY 3784 GCGCTTTTGTCTGGCTACCACTTTCCCGAGAGCAGGCCCCCGAGCCCCCTTCAGGC 3843
DB 3785 GCGCTTTTGTCTGGCTACCACTTTCCCGAGAGCAGGCCCCCGAGCCCCCTTCAGGC 3844
QY 3844 CCAGACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTGAT 3903
DB 3845 CCAGACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTGAT 3904
QY 3904 GCAGATATGCTGACAGGAACAGTCTGTGGATGACATGATCAGTGTCTAAGGAAGCAG 3963

QY 662 AGAAACGAGAGAGAGCTCAAAAGTCCCTGGCTCATGACAGGAGTGGCTTGGCAAAAC 721
DB 713 AGAAACGAGAGAGAGCTCAAAAGTCCCTGGCTCATGACAGGAGTGGCTTGGCAAAAC 772
QY 722 CCTCCCGAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTCCAGGAGGATGAGTCTC 781
DB 773 CCTCCCGAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTCCAGGAGGATGAGTCTC 832
QY 782 CACTCGGCGCCCATATGTTAGAAAACCCCGAGTTTCAACAGGCTCTGAGGAACCCAG 841
DB 833 CACTCGGCGCCCATATGTTAGAAAACCCCGAGTTTCAACAGGCTCTGAGGAACCCAG 892
QY 842 GCCTTGGGCACTCTGTTTAAAGACGCTTGGCAGGCGCTACGGCCGCTCTGCCCTCGAT 901
DB 893 GCCTTGGGCACTCTGTTTAAAGACGCTTGGCAGGCGCTACGGCCGCTCTGCCCTCGAT 952
QY 902 CAGAACTCCACAACTGATCAGCCCTTGAACCTGCAATGCTGAACACAGCTGGAACCTGCACC 961
DB 953 CAGAACTCCACAACTGATCAGCCCTTGAACCTGCAATGCTGAACACAGCTGGAACCTGCACC 1012
QY 962 ACCCCAGGACGAGGCGCCCTTGGCCCTTGCACACGACCCCTTCCCTTATAGCAGACTGC 1021
DB 1013 ACCCCAGGACGAGGCGCCCTTGGCCCTTGCACACGACCCCTTCCCTTATAGCAGACTGC 1072
QY 1022 CTCTTCCCTTCCCATTCACCTCTCCAGCCCTGGAACCTCAACCTCTGAGAGTCTTCC 1081
DB 1073 CTCTTCCCTTCCCATTCACCTCTCCAGCCCTGGAACCTCAACCTCTGAGAGTCTTCC 1132
QY 1082 TGGGCAAACTGCGCTGTGAGACGCGAGAACCTTGCCTGACCCACACCTGAGCAAAAC 1141
DB 1133 TGGGCAAACTGCGCTGTGAGACGCGAGAACCTTGCCTGACCCACACCTGAGCAAAAC 1192
QY 1142 TGGCCTGTGTAGACAGTCAAAAGCCCTTGGCCCTGACACCTGAGGACCCAGCTGCTGT 1201
DB 1193 TGGCCTGTGTAGACAGTCAAAAGCCCTTGGCCCTGACACCTGAGGACCCAGCTGCTGT 1252
QY 1202 CTCTGTGCTCCCATGAGAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGCA 1261
DB 1253 CTCTGTGCTCCCATGAGAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGCA 1312
QY 1262 GCGTGTGCTCAAGCAGGCGCCACAGCTGACCCAGCTGCGCAAGACCTGGGCGACACGG 1321
DB 1313 GCGTGTGCTCAAGCAGGCGCCACAGCTGACCCAGCTGCGCAAGACCTGGGCGACACGG 1372
QY 1322 GCTCCAGATCCCGGAGGCGCCCAAACTGAGGCAAACTGAGGCTGTCTGTCTCACTG 1381
DB 1373 GCTCCAGATCCCGGAGGCGCCCAAACTGAGGCAAACTGAGGCTGTCTGTCTCACTG 1432
QY 1382 AGAAACTCAAGCAGTGTATGATACCGAGAGAGTCCACTGGGCCACGCAACCAGC 1441
DB 1433 AGAAACTCAAGCAGTGTATGATACCGAGAGAGTCCACTGGGCCACGCAACCAGC 1492
QY 1442 TCCGCTGGGAGAGGCTCTTCGAGAGGTGCACAGGATGGAGGACAAAGCAGTGGCT 1501
DB 1493 TCCGCTGGGAGAGGCTCTTCGAGAGGTGCACAGGATGGAGGACAAAGCAGTGGCT 1552
QY 1502 TCCAGTGGCTCAAAAAGTGGCCCTGGAAGTATTTTCGGCAGGAGCTGATGGCAT 1561
DB 1553 TCCAGTGGCTCAAAAAGTGGCCCTGGAAGTATTTTCGGCAGGAGCTGATGGCAT 1612
QY 1562 GTGAGGATGTACCTCACCAGAAATGTCCTTTGTATGGAGCTGTGAGAGAGGGCTT 1621
DB 1613 GTGAGGATGTACCTCACCAGAAATGTCCTTTGTATGGAGCTGTGAGAGAGGGCTT 1672
QY 1622 GGTCAACATCTTATGAGCTGTGGAAGGTGGCTCCCTGGGCGCAGCTGTCAAGGAGC 1681
DB 1673 GGTCAACATCTTATGAGCTGTGGAAGGTGGCTCCCTGGGCGCAGCTGTCAAGGAGC 1732
QY 1682 AGGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGCCAGGCCCTGGAGGCTGTGG 1741
DB 1733 AGGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGCCAGGCCCTGGAGGCTGTGG 1792
QY 1742 AATACCTCCTCAGGAAGGATTTCTGCATGGGGAGCTCAAAAGCTGACAAACGCTGCTCTGT 1801

DB 1793 AATACCTCCTCAGGAAGGATTTGTCATGGGAGCGTCAAAAGCTGACAAACGCTGCTCTGT 1852
QY 1802 CCAGCATGGGAGCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTG 1861
DB 1853 CCAGCATGGGAGCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTG 1912
QY 1862 ATGCGCTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGCACAGAGACCCACATGG 1921
DB 1913 ATGCGCTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGCACAGAGACCCACATGG 1972
QY 1922 CTCCGAGGTGGTCTGGGACGAGCTGCGACGCCAAGGTGATGCTGAGAGAGCTGCT 1981
DB 1973 CTCCGAGGTGGTCTGGGACGAGCTGCGACGCCAAGGTGATGCTGAGAGAGCTGCT 2032
QY 1982 GTATGATGCTGCATGCTCAACGCTGCCACCCCTGACTCAGTTCTTCCGAGGCGCGC 2041
DB 2033 GTATGATGCTGCATGCTCAACGCTGCCACCCCTGACTCAGTTCTTCCGAGGCGCGC 2092
QY 2042 TCTGCTCAAGATTGCCAGGAGCTCCGCTGTGAGGAGATCCACCTCTCTCGGCC 2101
DB 2093 TCTGCTCAAGATTGCCAGGAGCTCCGCTGTGAGGAGATCCACCTCTCTCGGCC 2152
QY 2102 CTCTCAGCAGCCAGGCCATCCAGAGGGCTGAGGAAAGAGGCCATCCACCGCTGTCTG 2161
DB 2153 CTCTCAGCAGCCAGGCCATCCAGAGGGCTGAGGAAAGAGGCCATCCACCGCTGTCTG 2212
QY 2162 CAGCGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCC 2221
DB 2213 CAGCGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCC 2272
QY 2222 CTTGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACACC 2281
DB 2273 CTTGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACACC 2332
QY 2282 AGACCTCTCATGCGCCGAGAGAGCTTTTCGCAAGGGCCCGAGGGCCCGGCCAGCTG 2341
DB 2333 AGACCTCTCATGCGCCGAGAGAGCTTTTCGCAAGGGCCCGAGGGCCCGGCCAGCTG 2392
QY 2342 AGGAGACAAAGGAGAGGCCCTTAAGTCCAGCTCCTCTCCACAGAGCCCCCAGAGC 2401
DB 2393 AGGAGACAAAGGAGAGGCCCTTAAGTCCAGCTCCTCTCCACAGAGCCCCCAGAGC 2452
QY 2402 CAAAACAGTCTCTCTTGTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGAAACCTTTAC 2461
DB 2453 CAAAACAGTCTCTCTTGTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGAAACCTTTAC 2512
QY 2462 CTCTGTCTCTCTGGAGCGCCCTTGCAGAAACCCCGAGCTCACAGAGCGGAAAGCAA 2521
DB 2513 CTCTGTCTCTCTGGAGCGCCCTTGCAGAAACCCCGAGCTCACAGAGCGGAAAGCAA 2572
QY 2522 CCGTCCCGGAGCAGGAACTGCAGCAGCTGGAAATAGAAATATTCTCTGCTGCTCAGCATCCAGAGCTCT 2581
DB 2573 CCGTCCCGGAGCAGGAACTGCAGCAGCTGGAAATAGAAATATTCTCTCAACAGCTGTCTCC 2632
QY 2582 AGCCATTTCTCTGGAGGAGCAGGAAATTTCTCTGCTGCTCAGCATCCAGAGCTCT 2641
DB 2633 AGCCATTTCTCTGGAGGAGCAGGAGCAATTTCTCTGCTGCTCAGCATCCAGAGCTCT 2692
QY 2642 CCCTGTGGATGACAGTGAGAAAGAACCCATCAAGGGCTCTCAAGCTCTCGGGGACACCC 2701
DB 2693 CCCTGTGGATGACAGTGAGAAAGAACCCATCAAGGGCTCTCAAGCTCTCGGGGACACCC 2752
QY 2702 TGAGCTCAGGCGTACACTCTGGAGCAGCCAGCGGAGCTCGAGCTCCAGCTGGAAACA 2761
DB 2753 TGAGCTCAGGCGTACACTCTGGAGCAGCCAGCGGAGCTCGAGCTCCAGCTGGAAACA 2812
QY 2762 TGGTGTGGCCCGGGGCGGCCACCGACACCCCAAGCTATTTCATGCTGTGAAAGTCC 2821
DB 2813 TGGTGTGGCCCGGGGCGGCCACCGACACCCCAAGCTATTTCATGCTGTGAAAGTCC 2872
QY 2822 AATACAGTCTCTTAATGGTGAAACCTCTGCATCCGCGAGTTCCACCGGCTCAAGTGG 2881

Db	2873	AAATACAGTCTCTTAATAGTGGTGAACA	CTGCACATCCGGGATGTTCCACCGGGTCAAAAGTGG	29392
Qy	2882	GAGACATCGCCACTGGCATCAGCAGCCAGAT	CCCGAGTGCAGCTTCAGCTTTGGTCAACCA	2941
Db	2933	GAGACATCGCCACTGGCATCAGCAGCCAGAT	CCCGAGTGCAGCTTCAGCTTTGGTCAACCA	2992
Qy	2942	AAGACGGGACGCTTTGCTGCTACGACATGAGAGTGCAGAT	CTGGGGATCGACTGGAGT	3001
Db	2993	AAGACGGGACGCTTTGCTGCTACGACATGAGAGTGCAGAT	CTGGGGATCGACTGGAGT	3052
Qy	3002	GCACATGGCCCCGTGATGGGAGCTTCCCTTGGAGCTGGAGGCTCAAGCATGGCCAGCTGG	3061	
Db	3053	GCACTGGCCCCGTGATGGGAGCTTCCCTTGGAGCTGGAGGCTCAAGCATGGCCAGCTGG	3112	
Qy	3062	AGAACAGGCCCCATAACCTTGCCCTCCACCGCCGGCTCCACATGCGGAAAGCAGCTTCC	3121	
Db	3113	AGAACAGGCCCCATAACCTTGCCCTCCACCGCCGGCTCCACATGCGG	AGCAGCTTCC	3171
Qy	3122	TGCTCGGTGCAAGATGTGCGCTGAAACACAGGCTCAGCGGTTCCAGGGGATTTGCCAG	3181	
Db	3172	TGCTCGGTGCAAGATGTGCGCGCTG-AAACACAGGCTCAGCGGTTCCAGGGATTTGCCAG	3230	
Qy	3182	CCCCCGGCTCACAGTGGGAAACAGGCGCTCGCAGCAGCAGTGGGGGCAAGCAGATG	3241	
Db	3231	-CCCCCGGCTCAGAGTGGGAAACAGGCGCTCGCAGCAGCAGTGGGGGCAAGCAGATG	3289	
Qy	3242	CCTCCAGGATTTTACACACTGAGCGCTGCCACCCCTGCTGAAACAAACA-TCCGCCACAGT	3300	
Db	3290	CCTCCAGGATTTTACACACTGAGCGCTGCCACCCCTGCTGAGAAACACTCCGCCACAGT	3349	
Qy	3301	GAAGAGACAGAAGGAGATGGCAGGATTAACCTGGGAAACAAACACAGGATCTTTTCT	3360	
Db	3350	GAAGAGACAG-AGGAGATGGCAGGATTAACCTCGGAAACAAACAGG---ATCTTCTC	3404	
Qy	3361	GCCCTCGTCCAGTTCGAGTTGGCCCTGACCCGCTTGGATCAGTGACCAATTTGTTGGCAGAC	3420	
Db	3405	TGCCCTGCTCCAGTTCGAGTTGGCCCTGACCCGCTTGGATCAGTGACCAATTTGCTGGCAGAC	3464	
Qy	3421	AGGGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGGAGGCGCCCTTCGGGCCCTCAAC	3480	
Db	3465	AGGGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGGAGGCGCCCTTCGGGCCCTCAAC	3524	
Qy	3481	CTCCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGGCCCAAACTCAGGTTCAGTGCGAGAA	3540	
Db	3525	CTCCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGGCCCAAACTCAGGTTCAGTGCGAGAA	3584	
Qy	3541	CCAGGTGACAGAGGTATGCCCGCCGTTAGTTTAAGGGGGCCCTCTAAACCCCTTCGCTGGC	3600	
Db	3585	CCAGGTGACAGAGGTATGCCCGCCGTTAGTTTAAGGGGGCCCTCTAAACCCCTTCGCTGGC	3644	
Qy	3601	CTCACCTGGCCAGCTCACCCCTTTGGGTGTAGGGGAAAGATGCGCTGACCTGGGGAAG	3660	
Db	3645	CTCACCTGGCCAGCTCACCCCTTTGGGTGTAGGGGAAAGATGCGCTGACCTGGGGAAG	3704	
Qy	3661	GCTCCCTGGTAGAATACACCACTTTTCAGGTTTGTGCAACACAGGTCCTGAGTTGACC	3720	
Db	3705	GCTCCCTGGTAGAATACACCACTTTTCAGGTTTGTGCAACACAGGTCCTGAGTTGACC	3764	
Qy	3721	TCTGGTTGAGCCAAAGGACCAAGAGGTGTAAAGTGAAGTGTTCAGTCCCGCAGACA	3780	
Db	3765	TCTGGTTGAGCCAAAGGACCAAGAGGTGTAAAGTGAAGTGTTCAGTCCCGCAGACA	3824	
Qy	3781	TGTGCCCTTTGCTGCTGGCTACCACTTTCCCGAGAGCAGCGGCCCGCAGGCCCTTCA	3840	
Db	3825	TGTGCCCTTTGCTGCTGGCTACCACTTTCCCGAGAGCAGCGGCCCGCAGGCCCTTCA	3884	
Qy	3841	GGCCACAGCACTGCCCCAGATCTGCTGGCAGCTCAGTTCCCTCATCTGTAAAGGTGAAGGGT	3900	
Db	3885	GGCCACAGCACTGCCCCAGATCTGCTGGCAGCTCAGTTCCCTCATCTGTAAAGGTGAAGGGT	3944	
Qy	3901	GATCGAGATATGCTGCACAGGAACAGTCTGTGTGATGGACATGATCAGTCTAAGGAAAG	3960	
Db	3945	GATCGAGATATGCTGCACAGGAACAGTCTGTGTGATGGACATGATCAGTCTAAGGAAAG	4004	

QY	3961	CAGCAGAGAGAGAC-TCCGGCGCCCGACGCCACATATCAGTGTCCAGCGTGTGTTCC	4019
DB	4005	CAGCAGAGAGAGACGCTCCGGCGCCCGACGCCACATATCAGTGTCCAGCGTGTGTTCC	4064
QY	4020	CCAGAGCACAGCTCAGCATCACACTGACACTCACCCCTGCCCTGCCCTGCAGAGGGA	4079
DB	4065	CCAGAGCACAGCTCAGCATCACACTGACACTCACCTGCCCTGCCCTGCAGAGGGA	4124
QY	4080	CTGCCGACGGACATTGTGCACTCTGTATGACCTCAAAAGCACTTTTCAATGGCTGCCCTCTGGCA	4139
DB	4125	CTGCCGACGGACATTGTGCACTCTGTATGACCTCAAAAGCACTTTTCAATGGCTGCCCTCTGGCA	4184
QY	4140	GGGCAGGGCAGGGCAGTGACACTGTGTAGGACATAGCAAGCCAGGAGATGGGTGAAGGA	4199
DB	4185	GGGCAGGGCAGGGCAGTGAGCACTGTGTAGGACATAGCAAGCCAGGAGATGGGTGAAGGA	4244
QY	4200	CACAGTCTTGAGCTGTGCCACATGCATGTGACTCTCTCAAACCTCTTCCAGATTTCTCTAAG	4259
DB	4245	CACAGTCTTGAGCTGTGCCACATGCATGTGACTCTCTCAAACCTCTTCCAGATTTCTCTAAG	4304
QY	4260	AATAGCACCCCTTCCCATTTGCCCHAGCTTAGCCTTTCTCCAGGGGAGCTACTCAGG	4319
DB	4305	AATAGCACCCCTTCCCATTTGCCCHAGCTTAGCCTTTCTCCAGGGGAGCTACTCAGG	4364
QY	4320	ACTCACCTGACATTAATACAGCTGTGAATCGTCAGGGGGTGTCGTCTAGCCTCAACCTCC	4379
DB	4365	ACTCACCTGACATTAATACAGCTGTGAATCGTCAGGGGGTGTCGTCTAGCCTCAACCTCC	4424
QY	4380	TGGGGCAGGGGACGCCGAGACTCCGTGGGAGAACTCATTTCCACATCTTTGCCAAGACAG	4439
DB	4425	TGGGGCAGGGGACGCCGAGACTCCGTGGGAGAACTCATTTCCACATCTTTGCCAAGACAG	4484
QY	4440	CCTTTGTCCAGCTGTCCACATTGAGTCCAGACTGCTCTCCGGGGAGAGAGCCCCGGCCCCCA	4499
DB	4485	CCTTTGTCCAGCTGTCCACATTGAGTCCAGACTGCTCTCCGGGGAGAGAGCCCCGGCCCCCA	4544
QY	4500	GCACATAAAGAACTGCAGCCTTGGTACTCCAGAGTCTGGGTCTGTAGAGAACTCTTTGTAA	4559
DB	4545	GCACATAAAGAACTGCAGCCTTGGTACTCCAGAGTCTGGGTCTGTAGAGAACTCTTTGTAA	4604
QY	4560	GCAATAAAGTTTGGGGTGATGACAAATGTT	4589
DB	4605	GCAATAAAGTTTGGGGTGATGACAAATGTT	4634

RESULT 7
AAV73917
ID AAV73917 standard: CDNA: 3156 BP.

AAV73917:

04-MAR-1999 (first entry)

Human NIK cDNA.

KW NIK; Nuclear factor-kappa B; NF-kB; signal transduction; TNF; human;
 KW tumour necrosis factor; NF-kB-inducing kinase; screening; interaction;
 KW modulator; immune response; inflammatory response; viral gene;
 KW diagnostic; therapy; ds.

XX Homo sapiens.

XX PN US5854003-A.

XX
PD
29-DEC-1998.XX
PF 26-FEB-1998: 98US-00032475.

XX 03-JUL-1997: 97US-00887518.

XX
PA (TULA-) TULARIK INC.

PI Wu L, Rothe M;
XX WPI; 1999-094902/08.
XX
PT Screening agents for modulating interaction of nuclear factor kappaB
PT inducing kinase - with kinase-binding target, useful for controlling
PT levels of the kinase, for treatment and diagnosis of conditions
PT associated with e.g. inhibition of signal transduction by tumour necrosis
XX factor.
XX
PS Disclosure; Col 11-14; 16pp; English.
XX
CC This sequence encodes a novel human nuclear factor-kappaB (NFkB) -inducing
CC kinase (NIK) which is used in a method for screening for agents that
CC modulate the interaction of NIK with a NIK-binding target. The encoded
CC protein can be used as a modulator of cellular functions at the NIK
CC level, or for development of such compounds. NFkB is involved in
CC expression of many immune and inflammatory responses and of some
CC important viral genes. The protein may be used diagnostically and
CC therapeutically, in conditions associated with abnormal utilisation of
CC pathways that involve NFkB, e.g. inhibition of signal transduction by
CC tumour necrosis factor (TNF)
XX
SQ Sequence 3156 BP; 736 A; 991 C; 895 G; 534 T; 0 U; 0 Other;
Query Match 68.5%; Score 3149.6; DB 2; Length 3156;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 232 ATGGCAGTGATGGAATGCTGCCAGGTGCCCCCTGCTCAGCAGTGGGGCAGCAGAAAG 291
DB 1 ATGGCAGTGATGGAATGCTGCCAGGTGCCCCCTGCTCAGCAGTGGGGCAGCAGAAAG 60
QY 292 GAACTCCCAAGCCAAAG 351
DB 61 GAACTCCCAAGCCAAAG 120
QY 352 AAGCTTGAAGCCGTGGAG 411
DB 121 AAGCTTGAAGCCGTGGAG 180
QY 412 GTGATTACAAAGGACAGCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 471
DB 181 GTGATTACAAAGGACAGCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 472 ATCCGCCAGGCTGAGTGAG 531
DB 241 ATCCGCCAGGCTGAGTGAG 300
QY 532 TTGATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAT 591
DB 301 TTGATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAT 360
QY 592 GTGGCCCATGTACAGAGGCAAAATATGCCCCGTGTGTGGAGAGGAAAGCGTCGAGC 651
DB 361 GTGGCCCATGTACAGAGGCAAAATATGCCCCGTGTGTGGAGAGGAAAGCGTCGAGC 420
QY 652 AAAGCCCGAAGAAACGGAAG 711
DB 421 AAAGCCCGAAGAAACGGAAG 480
QY 712 TTGGCCAAACCCCTCCCAAG 771
DB 481 TTGGCCAAACCCCTCCCAAG 540
QY 772 GATGAGTCTCACTCGGCGCCCAATATGTTAGAAACACCCCGCAGTTTACCAAGCCCTCTG 831
DB 541 GATGAGTCTCACTCGGCGCCCAATATGTTAGAAACACCCCGCAGTTTACCAAGCCCTCTG 600
QY 832 AAGGAAACAGAGCCCTTGGCAACTCTGTTTAAAGCAGCTTGGCAGAGGCGCTACGCGCGCT 891
DB 601 AAGGAAACAGAGCCCTTGGCAACTCTGTTTAAAGCAGCTTGGCAGAGGCGCTACGCGCGCT 660

QY 892 CTGCCTCGATCAGAACTCCCAAACTGATCAGCCCCCTTGCATATGTGTGAACCAACGCTGTGG 951
DB 661 CTGCCTCGATCAGAACTCCCAAACTGATCAGCCCCCTTGCATATGTGTGAACCAACGCTGTGG 720
QY 952 AAACCTGCACCAACCCCAAGAGCGGAGGCCCCCTGCCCCCTGCCACCAAGCCCTTCCCCTAT 1011
DB 721 AAATGTCACCAACCCCAAGAGCGGAGGCCCCCTGCCCCCTGCCACCAAGCCCTTCCCCTAT 780
QY 1012 AGCAGAGCTGCTCATCTCCCTTCCCATTCACCTCTCCAGCCCTCGGAAACCTCACCTCTG 1071
DB 781 AGCAGAGCTGCTCATCTCCCTTCCCATTCACCTCTCCAGCCCTCGGAAACCTCACCTCTG 840
QY 1072 GAGTCTCTCTGGGCAAACTGGGCTGTGTAGACAGCCAGCAAAACCTTGGCTTGACCCACAC 1131
DB 841 GAGTCTCTCTGGGCAAACTGGGCTGTGTAGACAGCCAGCAAAACCTTGGCTTGACCCACAC 900
QY 1132 CTGAGCAAACTGCGCTGTGTAGACAGTCCAAAGCCCTGCTGCCACCAACCTCGAGGCC 1191
DB 901 CTGAGCAAACTGCGCTGTGTAGACAGTCCAAAGCCCTGCTGCCACCAACCTCGAGGCC 960
QY 1192 AGCTGCTGTCTGCTGTGCTGCCCATGAGAAGTCTTGTGGAGGAATACCTAGTGCATGCT 1251
DB 961 AGCTGCTGTCTGCTGTGCTGCCCATGAGAAGTCTTGTGGAGGAATACCTAGTGCATGCT 1020
QY 1252 CTGCAAGGCGCTGAGCTCAAGCCAGGCCACAGCTGACCCAGCTGGCCCAAGACCTGG 1311
DB 1021 CTGCAAGGCGCTGAGCTCAAGCCAGGCCACAGCTGACCCAGCTGGCCCAAGACCTGG 1080
QY 1312 GCAGCAGGCGCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACACGAGGGTGTCT 1371
DB 1081 GCAGCAGGCGCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACACGAGGGTGTCT 1140
QY 1372 CTGCTCACTGAGAACTCAAGCCAGTGGATTTAGTATACCGAGAAAGTCCACTTGGGCC 1431
DB 1141 CTGCTCACTGAGAACTCAAGCCAGTGGATTTAGTATACCGAGAAAGTCCACTTGGGCC 1200
QY 1432 ACGCACAGCTCCGCTGGGCGAGAGGCTCTTCCGAGAGGTGTCACAGATGGAGACAAG 1491
DB 1201 ACGCACAGCTCCGCTGGGCGAGAGGCTCTTCCGAGAGGTGTCACAGATGGAGACAAG 1260
QY 1492 CAGACTGGCTTCCAGTCCGCTGTCAAAAGGTGCGCTGGAAGTATTTTCGGGCGAGGAG 1551
DB 1261 CAGACTGGCTTCCAGTCCGCTGTCAAAAGGTGCGCTGGAAGTATTTTCGGGCGAGGAG 1320
QY 1552 CTGATGCGCATGTGCGAGGATTTGACCTCACAGAAATGCTCCCTTTGTATGAGAGCTGTGAGA 1611
DB 1321 CTGATGCGCATGTGCGAGGATTTGACCTCACAGAAATGCTCCCTTTGTATGAGAGCTGTGAGA 1380
QY 1612 GAAGGGCTTTGGGTCAAACATCTTTCATGAGAGCTGCTGGAAGGTGGCTCCCTGGGCCAGCTG 1671
DB 1381 GAAGGGCTTTGGGTCAAACATCTTTCATGAGAGCTGCTGGAAGGTGGCTCCCTGGGCCAGCTG 1440
QY 1672 GTCAAGAGACAGGGCTGTCTCCAGAGAGACCGGGCCCTGTACTACTTGGGCGAGGCCCTG 1731
DB 1441 GTCAAGAGACAGGGCTGTCTCCAGAGAGACCGGGCCCTGTACTACTTGGGCGAGGCCCTG 1500
QY 1732 GAGGCTCTGGAATACCTCCACTCACAGAGGATTTCTGCATGGGGACGTCGAAAGCTGACAAAC 1791
DB 1501 GAGGCTCTGGAATACCTCCACTCACAGAGGATTTCTGCATGGGGACGTCGAAAGCTGACAAAC 1560
QY 1792 GTGCTCTCTCCAGCAGTGGGAGCCACGAGCCCTCTGTGATCTTTGGCCCATGCTGTGTGT 1851
DB 1561 GTGCTCTCTCCAGCAGTGGGAGCCACGAGCCCTCTGTGACTTTTGGCCCATGCTGTGTGT 1620
QY 1852 CTTCAACCTGATGGCTGGGAAAGTCTTCTGCTACAGGGAGACTACATCCCTTGGGCGACAGAG 1911
DB 1621 CTTCAACCTGATGGCTGGGAAAGTCTTCTGCTACAGGGAGACTACATCCCTTGGGCGACAGAG 1680
QY 1912 ACCCACATGCTCCGAGAGTGTGCTCGGCGAGGAGCTGCGACCGCAAGAGTGGATGTCTGG 1971
DB 1681 ACCCACATGCTCCGAGAGTGTGCTCGGCGAGGAGCTGCGACCGCAAGAGTGGATGTCTGG 1740
QY 1972 AGCAGTGTCTATGATGTGCAATGCTCAAAGGCTGCCACCCCTGGACTCAGTTCTTTC 2031

1741 AGCAGCTGCTGATGATGCTGACATGCTCAAGGCTGCACCCCTGACTCAGTTCTTC 1800
2032 CGAGGGCCGCTCTGCTCAAGATGCCAGAGCCCTCCGCTGTGAGGGAGATCCACCC 2091
1801 CGAGGGCCGCTCTGCTCAAGATGCCAGAGCCCTCCGCTGTGAGGGAGATCCACCC 1860
2092 TCCTGGGCCCCCTCTCACAGCCAGGCCATCCAGAGGGGCTGAGAAAGAGCCATCCAC 2151
1861 TCCTGGGCCCCCTCTCACAGCCAGGCCATCCAGAGGGGCTGAGAAAGAGCCATCCAC 1920
2152 CGCTGTCTGACCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGT 2211
1921 CGCTGTCTGACCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGT 1980
2212 CTGAAGAGCCCTTGAGGGGAGAAATATAAGAAACAAGACATCCACCGCCAAATCAAGCC 2271
1981 CTGAAGAGCCCTTGAGGGGAGAAATATAAGAAACAAGACATCCACCGCCAAATCAAGCC 2040
2272 AATTACCACAGACCTTCATGCCAGCCGAGAGAGCTTCGCCAAGGGCCCCAGGGCCC 2331
2041 AATTACCACAGACCTTCATGCCAGCCGAGAGAGCTTCGCCAAGGGCCCCAGGGCCC 2100
2332 CGGCCAGCTGAGAGACAAAGCAGAGCCCTTAAGCTCCAGCCTCCTCTCCACAGAG 2391
2101 CGGCCAGCTGAGAGACAAAGCAGAGCCCTTAAGCTCCAGCCTCCTCTCCACAGAG 2160
2392 CCCCCAGAGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGAGAGAGTCTGGAGTGTGG 2451
2161 CCCCCAGAGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGAGAGAGTCTGGAGTGTGG 2220
2452 GAACCTTACCTTGTCTCTCCCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACCAGAG 2511
2221 GAACCTTACCTTGTCTCTCCCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACCAGAG 2280
2512 CGMAAGCAACCGTCCCGGAGCAGGAACCTGCAGCAGCTGGAATATTAATTCCTCAAC 2571
2281 CGMAAGCAACCGTCCCGGAGCAGGAACCTGCAGCAGCTGGAATATTAATTCCTCAAC 2340
2572 AGCTCTCCAGCAATTTCTCTGGAGAGCAGAGCAAAATTTCTCTGTGCTCAGCATC 2631
2341 AGCTCTCCAGCAATTTCTCTGGAGAGCAGAGCAAAATTTCTCTGTGCTCAGCATC 2400
2632 GACAGCTCTCCCTGCTGAGTACAGTGAGAGAACCCATCAAGGCTCTCAAGACTCG 2691
2401 GACAGCTCTCCCTGCTGAGTACAGTGAGAGAACCCATCAAGGCTCTCAAGACTCG 2460
2692 CGGGACACCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGCTCGAAGCTCC 2751
2461 CGGGACACCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGCTCGAAGCTCC 2520
2752 AGCTGGAACATGCTGTGGCCCGGGGGCCGCCACCGACACCCCAAGCTATTTCATGCT 2811
2521 AGCTGGAACATGCTGTGGCCCGGGGGCCGCCACCGACACCCCAAGCTATTTCATGCT 2580
2812 GTCAAGTCCAAATACAGTCTCTTAATGTTGACACTGCACATCCGGAGTTCCACCGG 2871
2581 GTCAAGTCCAAATACAGTCTCTTAATGTTGACACTGCACATCCGGAGTTCCACCGG 2640
2872 GTCAAGTGGGAGACATCGCACTTGGCATCAGCAGCCAGATCCAGCTCAGCCTTCAGC 2931
2641 GTCAAGTGGGAGACATCGCACTTGGCATCAGCAGCCAGATCCAGCTCAGCCTTCAGC 2700
2932 TTGGTCCAAAGACGGGAGCCTGTTGCTACGATGAGAGTGCAGACTCGGGCATC 2991
2701 TTGGTCCAAAGACGGGAGCCTGTTGCTACGATGAGAGTGCAGACTCGGGCATC 2760
2992 GACTCGAGTGACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGGTCAAGCAT 3051
2761 GACTCGAGTGACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGGTCAAGCAT 2820
3052 GGCCAGCTGGAGAACAGGGCCCTAACCTGCTCCACCGCGGCTCCACACTCCCGGAAA 3111

2821 GGCCAGCTGGAGAACAGGCCCTAAACCTGCTCCACCGCGGCTCCACACTGCGCGAAA 2880
3112 GCAGCCTTCTGCTCGGTGCACGATGCTGCCCTGAAACACAGGCTCAGCCGTTCCCGAG 3171
2881 GCAGCCTTCTGCTCGGTGCACGATGCTGCCCTGAAACACAGGCTCAGCCGTTCCCGAG 2940
3172 GGATTGCCAGCCCCCGGCTCACAGTGGGAACCAAGGCCCTCGCAGCAGAGGTGGGGC 3231
2941 GGATTGCCAGCCCCCGGCTCACAGTGGGAACCAAGGCCCTCGCAGCAGAGGTGGGGC 3000
3232 AAGCAGAAATGCTCCAGGATTTTACACCTGAGCCCTGCCCCACCTGCTGAAAACAT 3291
3001 AAGCAGAAATGCTCCAGGATTTTACACCTGAGCCCTGCCCCACCTGCTGAAAACAT 3060
3292 CCGCAGCTGAAGAGACAGAGGAGGATGGCAGAGTTACTTGGGGAACCAAAACAGGGA 3351
3061 CCGCAGCTGAAGAGACAGAGGAGGATGGCAGAGTTACTTGGGGAACCAAAACAGGGA 3120
3352 TCTTTTCTGCCCCCTGCTCCAGTCCAGTTCGCTGGA 3387
3121 TCTTTTCTGCCCCCTGCTCCAGTTCGCTGGA 3156

RESULT 8
AAV71603
ID AAV71603 standard; cDNA; 3156 BP.
XX
AC AAV71603;
XX
DT 04-FEB-1999 (first entry)
XX
DE Human NF-kB-inducing kinase (NIK) polypeptide encoding cDNA.
XX
KW NIK; IKB; NF-kB; TNF; kinase; nuclear factor kappa B; inhibition;
tumour necrosis factor; binding; genetic hybridisation; screening;
signal transduction; biopharmaceutical; human; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2844
FT /*tag= a
FT /product= "NIK polypeptide"
XX
XX US5843721-A.
XX
XX 01-DEC-1998.
XX
XX 03-JUL-1997; 97US-00887518.
XX
XX 03-JUL-1997; 97US-00887518.
XX
XX (TULA-) TULARIK INC.
XX
XX Wu L, Rothe M;
XX
XX WPI: 1999-044580/04.
XX
XX P-PSDB; AAW81564.
XX
XX Probe, vector or recombinant nucleic acid encoding a polypeptide, useful
especially human nuclear factor kappa-B-inducing kinase protein - useful
for producing recombinant protein.
XX
XX Claim 1; Col 11-14; 15pp; English.
XX
XX This cDNA encodes a nuclear factor kappa B (NF-kB)-inducing kinase (NIK)
polypeptide. The NIK polypeptide or its fragment has one or more
activities selected from kinase activity and inhibitory activity; IKB
kinase-alpha and beta binding activity and binding inhibitory activity;
tumour necrosis factor (TNF) receptor-associated factor 2 binding
activity and binding inhibitory activity; IKB binding activity and
binding inhibitory activity. NF-kB activating and inhibitory activity. A
vector containing the NIK nucleic acid can be used to transform host

CC cells for the recombinant production of the protein. The NIK nucleic acid
CC and the polypeptide may be used in diagnosis (e.g. genetic hybridisation
CC screen for NIK transcripsts), therapy (e.g. NIK kinase inhibitors to
CC inhibit TNF signal transduction), and in the biopharmaceutical industry
XX
SQ Sequence 3156 BP; 736 A; 991 C; 895 G; 534 T; 0 U; 0 Other;

Query Match 68.5%; Score 3149.6; DB 2; Length 3156;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy	232	ATGGCAGTGATGGAATGGCTGCGCCAGTGGCCCTGCTCAGCAGTGGGCGCAGAG	291
Db	1	ATGGCAGTGATGGAATGGCTGCGCCAGTGGCCCTGCTCAGCAGTGGGCGCAGAG	60
Qy	292	GAACTCCCAAGCAAGGAGAGACGGCCGACCTGGGGAGAAACACAGAGCTCCGCTTAC	351
Db	61	GAACTCCCAAGCAAGGAGAGACGGCCGACCTGGGGAGAAACACAGAGCTCCGCTTAC	120
Qy	352	AAGCTTGAGGCGGTGGAGAGCCCTGTGTTTCGGGAAAGTGGAGATCCTGAATGAC	411
Db	121	AAGCTTGAGGCGGTGGAGAGCCCTGTGTTTCGGGAAAGTGGAGATCCTGAATGAC	180
Qy	412	GTGATTACCAAGGGCAGACCAAGGAAGGCTCCGAGGCGAGGCGCAGCTCTCTATC	471
Db	181	GTGATTACCAAGGGCAGACCAAGGAAGGCTCCGAGGCGAGGCGCAGCTCTCTATC	240
Qy	472	ATCGCCAGGCTGAGTGTGAGAAAGCAAGAGTTTCAGCCACCTTTTCAGAGCGATT	531
Db	241	ATCGCCAGGCTGAGTGTGAGAAAGCAAGAGTTTCAGCCACCTTTTCAGAGCGATT	300
Qy	532	TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAT	591
Db	301	TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAT	360
Qy	592	GTGGCCCATGTACAGAGGCAAAATGCCCGTGTGTGGAGGGAAGAGCGTCGAGC	651
Db	361	GTGGCCCATGTACAGAGGCAAAATGCCCGTGTGTGGAGGGAAGAGCGTCGAGC	420
Qy	652	AAAGCCGGAAGAAACGGAAGAGAGCTCAAAAGTCCCTGGCTCATGCAAGAGTGGCC	711
Db	421	AAAGCCGGAAGAAACGGAAGAGAGCTCAAAAGTCCCTGGCTCATGCAAGAGTGGCC	480
Qy	712	TTGGCCAAACCCCTCCCGAGACCCCTGAGCAGGAGAGCTGCACCATCCCAAGTGCAGGAG	771
Db	481	TTGGCCAAACCCCTCCCGAGACCCCTGAGCAGGAGAGCTGCACCATCCCAAGTGCAGGAG	540
Qy	772	GATGAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCCTCTG	831
Db	541	GATGAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCCTCTG	600
Qy	832	AAGGAACAGGCTTGGGCAACTCTGTTTAAAGCAGCTTGGGAGGCGCTACGCGCGCT	891
Db	601	AAGGAACAGGCTTGGGCAACTCTGTTTAAAGCAGCTTGGGAGGCGCTACGCGCGCT	660
Qy	892	CTGCCTCGATCAGAACTCCAAACTGATCAGCCCTTGCATATGTTGAAACACAGCTGTGG	951
Db	661	CTGCCTCGATCAGAACTCCAAACTGATCAGCCCTTGCATATGTTGAAACACAGCTGTGG	720
Qy	952	AAATGCAACCAACCCAGGAGCGGCGCCCTGCGCCCTGCGCCACGACCCCTTCCCTTAT	1011
Db	721	AAATGCAACCAACCCAGGAGCGGCGCCCTGCGCCCTGCGCCACGACCCCTTCCCTTAT	780
Qy	1012	AGCAGATGCTCATCTCCCTTCCCATTCACCCCTCTCCAGGCTTGGAAACCTCAGCCTCTG	1071
Db	781	AGCAGATGCTCATCTCCCTTCCCATTCACCCCTCTCCAGGCTTGGAAACCTCAGCCTCTG	840
Qy	1072	GAGTCTTCTCGGCAAACTGCGCTGTGTAGACAGCCAGAAACCTTGCCTGACCCACAC	1131
Db	841	GAGTCTTCTCGGCAAACTGCGCTGTGTAGACAGCCAGAAACCTTGCCTGACCCACAC	900
Qy	1132	CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGGCC	1191

Db	901	CTGAGCAAACTGGGCTGTGTGTAGACAGTCCAAAGGCCCTGCTGCGCCACACCTGGAGGCC	960
Qy	1192	AGCTGCTCTGCTCGTGTGGTCCCATGAGAAATTTTCTGTGGAGGAATACCTAGTGCATGCT	1251
Db	961	AGCTGCTCTGCTCGTGTGGTCCCATGAGAAATTTTCTGTGGAGGAATACCTAGTGCATGCT	1020
Qy	1252	CTGCAAGGACGCTGTAGCTCAAGCCAGGCCCAACAGCTTGACAGCCTGGCCAAAGCCTGG	1311
Db	1021	CTGCAAGGACGCTGTAGCTCAAGCCAGGCCCAACAGCTTGACAGCCTGGCCAAAGCCTGG	1080
Qy	1312	GCAGCAGGGGCTCCAGATCCCGGAGGCCAGCCCAAAAATCTGAGGACAAAGAGGCTGTC	1371
Db	1081	GCAGCAGGGGCTCCAGATCCCGGAGGCCAGCCCAAAAATCTGAGGACAAAGAGGCTGTC	1140
Qy	1372	CTGCTCACTGAGAACTCAAGCCAGTGGATTTATGAGTACCCGAGAAAGTCCACTGGGCC	1431
Db	1141	CTGCTCACTGAGAACTCAAGCCAGTGGATTTATGAGTACCCGAGAAAGTCCACTGGGCC	1200
Qy	1432	ACGCACAGCTCCGCTGGGCGAGAGGCTCTTCGAGAGAGGTGCACAGATGGAGGACAAG	1491
Db	1201	ACGCACAGCTCCGCTGGGCGAGAGGCTCTTCGAGAGAGGTGCACAGATGGAGGACAAG	1260
Qy	1492	CAGACTGGCTTCCAGTGGCTGTCAAAAAGGTGGCCCTGGAAGTATTTTCGGGCGAGAG	1551
Db	1261	CAGACTGGCTTCCAGTGGCTGTCAAAAAGGTGGCCCTGGAAGTATTTTCGGGCGAGAG	1320
Qy	1552	CTGATGCGATGTGAGGATGACCTCAACCCAGAAATTTGCTCCCTTTGTATGGAGCTGTGAGA	1611
Db	1321	CTGATGCGATGTGAGGATGACCTCAACCCAGAAATTTGCTCCCTTTGTATGGAGCTGTGAGA	1380
Qy	1612	GAAGGGCTTTGGGTCAAACATCTTCATGAGCTGTCTGGAAGGTGGCTCCCTGGGCCAGCTG	1671
Db	1381	GAAGGGCTTTGGGTCAAACATCTTCATGAGCTGTCTGGAAGGTGGCTCCCTGGGCCAGCTG	1440
Qy	1672	GTCAAGGACGAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCGAGGCCCTG	1731
Db	1441	GTCAAGGACGAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCGAGGCCCTG	1500
Qy	1732	GAGGGTCTGGAATACCTCCACTCAAGAGGATCTGATGGGGAGCTCAAAAGCTGACAAAC	1791
Db	1501	GAGGGTCTGGAATACCTCCACTCAAGAGGATCTGATGGGGAGCTCAAAAGCTGACAAAC	1560
Qy	1792	GTGCTCTCTGTCCAGCGATGGGAGCCACGACCCCTCTGTGTGATTTTGGCCATGCTGTGT	1851
Db	1561	GTGCTCTCTGTCCAGCGATGGGAGCCACGACCCCTCTGTGTGATTTTGGCCATGCTGTGT	1620
Qy	1852	CTTCAACCTGATGCTGGGAAAGTCTTGTCTACAGGGGAGCTACATCCCTGGGCGACAG	1911
Db	1621	CTTCAACCTGATGCTGGGAAAGTCTTGTCTACAGGGGAGCTACATCCCTGGGCGACAG	1680
Qy	1912	ACCCACATGGCTCCGAGGTGGTGTCTGGGAGGAGCTGCGACGCAAGGTGGATGTCTGG	1971
Db	1681	ACCCACATGGCTCCGAGGTGGTGTCTGGGAGGAGCTGCGACGCAAGGTGGATGTCTGG	1740
Qy	1972	AGCAGCTGCTGTATGATGTGCAATGCTCAACGGCTGGCCAGCCCTTGAGCTCAGTTCCTC	2031
Db	1741	AGCAGCTGCTGTATGATGTGCAATGCTCAACGGCTGGCCAGCCCTTGAGCTCAGTTCCTC	1800
Qy	2032	CGAGGGCGCTCTGCTCTCAAGATTGCGAGGAGCTCCGCTGTGAGGAGATCCACCC	2091
Db	1801	CGAGGGCGCTCTGCTCTCAAGATTGCGAGGAGCTCCGCTGTGAGGAGATCCACCC	1860
Qy	2092	TCCTGCGCCCTCTCAGCGCCAGGCCATCCAGAGGGGCTGAGGAAAGGCCATCCAC	2151
Db	1861	TCCTGCGCCCTCTCAGCGCCAGGCCATCCAGAGGGGCTGAGGAAAGGCCATCCAC	1920
Qy	2152	CGCGTGTCTGCGAGGAGCTGGGAGGGAAGGTGAACCGGGCACTTACAGCAAGTGGAGGT	2211
Db	1921	CGCGTGTCTGCGAGGAGCTGGGAGGGAAGGTGAACCGGGCACTTACAGCAAGTGGAGGT	1980
Qy	2212	CTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCC	2271
Db	1981	CTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCC	2040

QY 2272 AATTACACAGACCTCATGCCAGCGGAGAGAGCTTTCCGCAAGGGCCCCAGGGCCC 2331
DB 2041 AATTACACAGACCTCATGCCAGCGGAGAGAGCTTTCCGCAAGGGCCCCAGGGCCC 2100
QY 2332 CGGCAGCTGAGAGACAACAGCGAGAGCCCTTAAGCTCCAGCCTCCTCTCCACAGAG 2391
DB 2101 CGGCAGCTGAGAGACAACAGCGAGAGCCCTTAAGCTCCAGCCTCCTCTCCACAGAG 2160
QY 2392 CCCCAGAGCCAAAAGTCTCTCCCTTGACTTTGAGCAAGAGAGTCTGGGATGTGG 2451
DB 2161 CCCCAGAGCCAAAAGTCTCTCCCTTGACTTTGAGCAAGAGAGTCTGGGATGTGG 2220
QY 2452 GAACCTTACCTCTGTCTCCCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACCAGAG 2511
DB 2221 GAACCTTACCTCTGTCTCCCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACCAGAG 2280
QY 2512 CGGAAGCAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAATAGAATTATTCCTCAAC 2571
DB 2281 CGGAAGCAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAATAGAATTATTCCTCAAC 2340
QY 2572 AGCCTGTCCAGCCATTTTCTCTGAGAGAGAGGAGCAAAATTTCTCTGTGCTCAGCATC 2631
DB 2341 AGCCTGTCCAGCCATTTTCTCTGAGAGAGAGGAGCAAAATTTCTCTGTGCTCAGCATC 2400
QY 2632 GACAGCTCTCCCTGTCCGATGACAGTGAGAGAACCCATCAAGGCCCTCTCAAGCTCG 2691
DB 2401 GACAGCTCTCCCTGTCCGATGACAGTGAGAGAACCCATCAAGGCCCTCTCAAGCTCG 2460
QY 2692 CGGACACCTCAGCTCAGGCGTACACTCTGGAGCAGCGCCGAGGCTCGAAGCTCC 2751
DB 2461 CGGACACCTCAGCTCAGGCGTACACTCTGGAGCAGCGCCGAGGCTCGAAGCTCC 2520
QY 2752 AGCTGAAACATGCTGTGCTCCCGGGGGCGGCCACCGAACAGCTATTTCATGCT 2811
DB 2521 AGCTGAAACATGCTGTGCTCCCGGGGGCGGCCACCGAACAGCTATTTCATGCT 2580
QY 2812 GTGAAGTCAATACAGTCTTAAATGTTGACACCTGCACATCCGGAGTTCCACCGG 2871
DB 2581 GTGAAGTCAATACAGTCTTAAATGTTGACACCTGCACATCCGGAGTTCCACCGG 2640
QY 2872 GTCAAAAGTGGAGACATCGCCACTGCGCATCAGCAGCCAGATCCAGCTCAGCCTTCAGC 2931
DB 2641 GTCAAAAGTGGAGACATCGCCACTGCGCATCAGCAGCCAGATCCAGCTCAGCCTTCAGC 2700
QY 2932 TTGGTCACAAAGACGGGAGCCTGTTGCTAGCAATGGAGGTGCCAGCTCGGGCATC 2991
DB 2701 TTGGTCACAAAGACGGGAGCCTGTTGCTAGCAATGGAGGTGCCAGCTCGGGCATC 2760
QY 2992 GACCTGAGTGACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTTCAAGCAT 3051
DB 2761 GACCTGAGTGACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTTCAAGCAT 2820
QY 3052 GGCAGCTGGAGAACAGGCGCTTAACCTGCTCCCTCCACCGCGCTCCACACTGCCGAAA 3111
DB 2821 GGCAGCTGGAGAACAGGCGCTTAACCTGCTCCCTCCACCGCGCTCCACACTGCCGAAA 2880
QY 3112 GCAGCCTTCTGCTCGGTGACGATGCTGCCTTGAAAAACAGGCTCAGCGTTCCAGG 3171
DB 2881 GCAGCCTTCTGCTCGGTGACGATGCTGCCTTGAAAAACAGGCTCAGCGTTCCAGG 2940
QY 3172 GGAATTGCAGCCCCCGGCTCAGTGGGNAACAGGGCTTCGAGCAGCAAGTGGGGC 3231
DB 2941 GGAATTGCAGCCCCCGGCTCAGTGGGNAACAGGGCTTCGAGCAGCAAGTGGGGC 3000
QY 3232 AAGCAGATTCCTCCAGGATTTTCACTCTGAGCCCTGCCCCACCTCTGTGAAAAACAT 3291
DB 3001 AAGCAGATTCCTCCAGGATTTTCACTCTGAGCCCTGCCCCACCTCTGTGAAAAACAT 3060
QY 3292 CCGCCACGTGAAGAGACAGAGGAGATGGCAGGATTTACCTGGGGAACAAACAGGGA 3351
DB 3061 CCGCCACGTGAAGAGACAGAGGAGATGGCAGGATTTACCTGGGGAACAAACAGGGA 3120

QY 3352 TCTTTTCTGCCCCCTGCTCCAGTCAGTTCAGTTCGCTTGA 3387
DB 3121 TCTTTTCTGCCCCCTGCTCCAGTCAGTTCAGTTCGCTTGA 3156
RESULT 9
AAV69285
ID AAV69285 standard; cDNA; 3156 BP.
XX AAV69285;
AC AAV69285;
XX 04-FEB-1999 (first entry)
XX Human NF-kB-inducing kinase (NIK) polypeptide encoding cDNA.
DE NIK; Ikb; NF-kB; TNF; kinase; nuclear factor kappa B; inhibition;
KW tumour necrosis factor; binding; genetic hybridisation; screening;
KW signal transduction; biopharmaceutical; immunogen; pharmacological;
KW transcription regulator; human; ds.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 1. .2844
FT /tag= a
FT /product= "NIK polypeptide"
XX
XX US944073-A.
XX 01-DEC-1998.
XX 13-FEB-1998; 98US-00023321.
XX 03-JUL-1997; 97US-00887518.
XX (TULA-) TULARIK INC.
XX Wu L, Rothe M;
XX WPI; 1999-044664/04.
XX P-PSDB; AAW81561.
XX New isolated peptide comprising a specified 947 amino acid sequence - has
PT e.g. kinase activity, kinase inhibitory activity, Ikb kinase-alpha
PT binding activity, and Ikb kinase-alpha binding inhibitory activity.
XX Disclosure; Col 11-14; 15pp; English.
XX This cDNA encodes a nuclear factor kappa B (NF-kB)-inducing kinase (NIK)
CC polypeptide. The NIK polypeptide or its fragment has one or more
CC activities selected from kinase activity and inhibitory activity; Ikb
CC kinase-alpha and beta binding activity and binding inhibitory activity;
CC tumour necrosis factor (TNF) receptor-associated factor 2 binding
CC activity and binding inhibitory activity; Ikb binding activity and
CC binding inhibitory activity, NF-kB activating and inhibitory activity.
CC The NIK nucleic acid and the polypeptide may be used in diagnosis (e.g.
CC genetic hybridisation screen for NIK transcripts), therapy (e.g. NIK
CC kinase inhibitors to inhibit TNF signal transduction), and in the
CC biopharmaceutical industry (e.g. as immunogens, reagents for isolating
CC other transcription regulators, and reagents for screening chemical
CC libraries for pharmacological agents)
XX
SQ Sequence 3156 BP; 736 A; 991 C; 895 G; 534 T; 0 U; 0 Other;
Query Match 68.5%; Score 3149.6; DB 2; Length 3156;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 232 ATGGCAGTGCATGAAATGGCTGCCAGTGCCTGGCTCAGCAGTGGGGCAGCAGAAG 291
DB 1 ATGGCAGTGCATGAAATGGCTGCCAGTGCCTGGCTCAGCAGTGGGGCAGCAGAAG 60
QY 292 GAACTCCCCCAAGCCAAAGGAGAGACGCCGCCACTTGGGGAAGAAACAGAGCTCCGCTTAC 351

Db 61 GAACTCCCCAAGCCAGGAGAGACCGCCACCTGGGGAGAAACAGAGCTCGTCTAC 120
Qy 352 AAGCTTGAGGCGGTGGAGAGAGCCCTGTGTTCTCGGAAAGTGGAGATCCTGAATGAC 411
Db 121 AAGCTTGAGGCGGTGGAGAGAGCCCTGTGTTCTCGGAAAGTGGAGATCCTGAATGAC 180
Qy 412 GTGATTACCAAGGGCAGACAGGAGAGAGCCCGAGGAGAGGCGGAGCTGCGATCTCTATC 471
Db 181 GTGATTACCAAGGGCAGACAGGAGAGAGCCCGAGGAGAGGCGGAGCTGCGATCTCTATC 240
Qy 472 ATCGCCCAAGGCTGAGTGTGAGAAATAGCAAGATTTCAGCCCACTTTTCAGAACGCAAT 531
Db 241 ATCGCCCAAGGCTGAGTGTGAGAAATAGCAAGATTTCAGCCCACTTTTCAGAACGCAAT 300
Qy 532 TTATCGCTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAT 591
Db 301 TTATCGCTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAT 360
Qy 592 GTGGCCCATGTACAGAGGGCAAAATGGCCGTGTGTGGAGAGGAAAGCTGCGAGC 651
Db 361 GTGGCCCATGTACAGAGGGCAAAATGGCCGTGTGTGGAGAGGAAAGCTGCGAGC 420
Qy 652 AAAGCCCGAAGAAAGCAAGAGAGCTCAAGTCCCTGGCTCATGCAAGAGTGGCC 711
Db 421 AAAGCCCGAAGAAAGCAAGAGAGCTCAAGTCCCTGGCTCATGCAAGAGTGGCC 480
Qy 712 TTGGCCAAACCCCTCCAGGAGCCCTGAGCAGAGAGCTGCACCATCCAGTGCAGGAG 771
Db 481 TTGGCCAAACCCCTCCAGGAGCCCTGAGCAGAGAGCTGCACCATCCAGTGCAGGAG 540
Qy 772 GATGAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGAGTTCACCAAGCCTCTG 831
Db 541 GATGAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGAGTTCACCAAGCCTCTG 600
Qy 832 AAGGAACAGGCTTGGGCAACTCTGTTTAAAGCAGCTTGGGAGGGCTACGGCCGGCT 891
Db 601 AAGGAACAGGCTTGGGCAACTCTGTTTAAAGCAGCTTGGGAGGGCTACGGCCGGCT 660
Qy 892 CTGCTCGATCAGAACTCCAAACTGATCAGCCCTTTGCAATGCTGAAACCAAGTGGG 951
Db 661 CTGCTCGATCAGAACTCCAAACTGATCAGCCCTTTGCAATGCTGAAACCAAGTGGG 720
Qy 952 AAATGCAACACCCCGAGGAGGAGGCGCCCTGCGCCCTGCCAGCAGCCCTTCCCTAT 1011
Db 721 AAATGCAACACCCCGAGGAGGAGGCGCCCTGCGCCCTGCCAGCAGCCCTTCCCTAT 780
Qy 1012 AGCAGACTGCCTCATCCCTTCCCATTCACCCCTTCCAGCCCTGGAACCTCAGCCTCTG 1071
Db 781 AGCAGACTGCCTCATCCCTTCCCATTCACCCCTTCCAGCCCTGGAACCTCAGCCTCTG 840
Qy 1072 GAGTCCTTCTGGGCAAACTGGCCTGTGTAGACAGCCAGAAACCTTGCCTGAGCCACAC 1131
Db 841 GAGTCCTTCTGGGCAAACTGGCCTGTGTAGACAGCCAGAAACCTTGCCTGAGCCACAC 900
Qy 1132 CTGAGCAAACTGGCCTGTGTAGACAGTCCAAAGCCCTTGCCTGGCCCACTGAGGCC 1191
Db 901 CTGAGCAAACTGGCCTGTGTAGACAGTCCAAAGCCCTTGCCTGGCCCACTGAGGCC 960
Qy 1192 AGCTGCCTGTCTGTGGTGCCTAGAGAGTTCCTGTGGAGGAATCCTAGTGTGATGCT 1251
Db 961 AGCTGCCTGTCTGTGGTGCCTAGAGAGTTCCTGTGGAGGAATCCTAGTGTGATGCT 1020
Qy 1252 CTGAGGAGCGGTGAGCTCAAGCCAGGCGCCACAGCTGACAGCCTGGCCCAAGCCTGG 1311
Db 1021 CTGAGGAGCGGTGAGCTCAAGCCAGGCGCCACAGCTGACAGCCTGGCCCAAGCCTGG 1080
Qy 1312 CGAGCAGGGGCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAAGAGGTGTC 1371
Db 1081 CGAGCAGGGGCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAAGAGGTGTC 1140
Qy 1372 CTGCTCACTGAGAACTCAAGCCAGTGGATTATGAGTACCGAGAGAGTCCACTGGGCC 1431

Db 1141 CTGCTCACTGAGAACTCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCC 1200
Qy 1432 ACGCAACAGCTCCCGCTGGGAGAGGCTCTTTCGGAGAGGTGCACAGGATGGAGACAAG 1491
Db 1201 ACGCAACAGCTCCCGCTGGGAGAGGCTCTTTCGGAGAGGTGCACAGGATGGAGACAAG 1260
Qy 1492 CAGACTGGCTTCCAGTGGCTTCAAAAGGTGCGCCCTGGAGATATTTCCGGGAGAGGAG 1551
Db 1261 CAGACTGGCTTCCAGTGGCTTCAAAAGGTGCGCCCTGGAGATATTTCCGGGAGAGGAG 1320
Qy 1552 CTGATGGCATGTGAGGATTCAGCTCCAGAAATTTGCTCCCTTTGTATGGAGCTGTGAGA 1611
Db 1321 CTGATGGCATGTGAGGATTCAGCTCCAGAAATTTGCTCCCTTTGTATGGAGCTGTGAGA 1380
Qy 1612 GAAGGGCTTGGGTCAACATCTTCATGAGCTGTGGAAGGTGCTCCCTGGGCCAGCTG 1671
Db 1381 GAAGGGCTTGGGTCAACATCTTCATGAGCTGTGGAAGGTGCTCCCTGGGCCAGCTG 1440
Qy 1672 GTCAAGGAGCAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTG 1731
Db 1441 GTCAAGGAGCAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTG 1500
Qy 1732 GAGGTCTGGAATACCTCCACTCAGAAAGGATTCGATGGGAGCTCAAAAGCTGACAAC 1791
Db 1501 GAGGTCTGGAATACCTCCACTCAGAAAGGATTCGATGGGAGCTCAAAAGCTGACAAC 1560
Qy 1792 GTGCTCTGTCTCAGCGATGGGAGCCAGCGACCCCTCTGTGACTTTGGCCATGCTGTGT 1851
Db 1561 GTGCTCTGTCTCAGCGATGGGAGCCAGCGACCCCTCTGTGACTTTGGCCATGCTGTGT 1620
Qy 1852 CTTCAAACCTGATGGCTTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGGCACAGAG 1911
Db 1621 CTTCAAACCTGATGGCTTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGGCACAGAG 1680
Qy 1912 ACCCACATGCTCCGAGGTGTGCTGGGAGGAGTGCAGCGCCAGAGTGGATGTCTGG 1971
Db 1681 ACCCACATGCTCCGAGGTGTGCTGGGAGGAGTGCAGCGCCAGAGTGGATGTCTGG 1740
Qy 1972 AGCAGCTGCTGTATGATGTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTCTTTC 2031
Db 1741 AGCAGCTGCTGTATGATGTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTCTTTC 1800
Qy 2032 CGAGGGCGCTCTGCTCAAGATTGCCAGGAGCCTCCGCTGTGAGGGAGATCCACCC 2091
Db 1801 CGAGGGCGCTCTGCTCAAGATTGCCAGGAGCCTCCGCTGTGAGGGAGATCCACCC 1860
Qy 2092 TCCTGCGCCCTCTCAGCGCCAGGCCATCCAAAGGGGCTGAGGAAGAGCCCATCCAC 2151
Db 1861 TCCTGCGCCCTCTCAGCGCCAGGCCATCCAAAGGGGCTGAGGAAGAGCCCATCCAC 1920
Qy 2152 CGCTGTCTGCAGCGGAGCTGGGAGGGAAGGTGAAACGGGCACTTACAGCAAGTGGAGGT 2211
Db 1921 CGCTGTCTGCAGCGGAGCTGGGAGGGAAGGTGAAACGGGCACTTACAGCAAGTGGAGGT 1980
Qy 2212 CTGAAGAGCCCTTGGAGGGGAGAAATAAAGAACCAAGACATCCACCGCCAAATCAAGCC 2271
Db 1981 CTGAAGAGCCCTTGGAGGGGAGAAATAAAGAACCAAGACATCCACCGCCAAATCAAGCC 2040
Qy 2272 AATTACCACAGACCTTCCATGCCAGCGAGAGCTTTGCCCAAGGGCCCGAGGGCCC 2331
Db 2041 AATTACCACAGACCTTCCATGCCAGCGAGAGCTTTGCCCAAGGGCCCGAGGGCCC 2100
Qy 2332 CGGCCAGCTGAGGAGACACAGGAGAGCCCTTAAGCTTCCAGCTCTCTCTCCACACAGAG 2391
Db 2101 CGGCCAGCTGAGGAGACACAGGAGAGCCCTTAAGCTTCCAGCTCTCTCTCCACACAGAG 2160
Qy 2392 CCCCCAGAGCCAAACAAGTCTCTCCCTTGAATTTGAGCAGAGGAGTCTGGGATGTGG 2451
Db 2161 CCCCCAGAGCCAAACAAGTCTCTCCCTTGAATTTGAGCAGAGGAGTCTGGGATGTGG 2220
Qy 2452 GAACCTTACTCTGTCTCTCCCTGGAGCCAGCCCTCTGCCAGAAAACCCAGCTCACCAGAG 2511
Db 2221 GAACCTTACTCTGTCTCTCCCTGGAGCCAGCCCTCTGCCAGAAAACCCAGCTCACCAGAG 2280

QY 2512 CGGAAGCAACGTCGCCGAGCAGGAACCTGACGAGCTGGAATAGAAATTATTCCTCAAC 2571
Db |||||
2281 CGGAAGCAACGTCGCCGAGCAGGAACCTGACGAGCTGGAATAGAAATTATTCCTCAAC 2340
QY 2572 AGCCTGTCCAGCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGCCTCAGCATC 2631
Db |||||
2341 AGCCTGTCCAGCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGCCTCAGCATC 2400
QY 2632 GACGCTCTCCCTGTCCGATGACAGTGAGAGAGACCCATCAAGCCCTCTCAAGACTCG 2691
Db |||||
2401 GACGCTCTCCCTGTCCGATGACAGTGAGAGAGACCCATCAAGCCCTCTCAAGACTCG 2460
QY 2692 CGGGACACCTGAGCTCAGGCGCTACACTCTGGAGCAGCGCGAGGCTCGAAGCTCC 2751
Db |||||
2461 CGGGACACCTGAGCTCAGGCGCTACACTCTGGAGCAGCGCGAGGCTCGAAGCTCC 2520
QY 2752 AGCTGGAAATGGTGTGCTGCCCGGGGGGGCCCAACGACACCCCAAGCTATTTCAATGCT 2811
Db |||||
2521 AGCTGGAAATGGTGTGCTGCCCGGGGGGGCCCAACGACACCCCAAGCTATTTCAATGCT 2580
QY 2812 GTGAAGTCCAAATACAGTCTTTAATGCTGAACCTGCACATCGGGAGTTCCACGG 2871
Db |||||
2581 GTGAAGTCCAAATACAGTCTTTAATGCTGAACCTGCACATCGGGAGTTCCACGG 2640
QY 2872 GTCAAAGTGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGC 2931
Db |||||
2641 GTCAAAGTGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGC 2700
QY 2932 TTGGTCAACAAAGACGGGAGCCTGTTCGCTACGATGAGAGTGCCAGACTCGGGCATC 2991
Db |||||
2701 TTGGTCAACAAAGACGGGAGCCTGTTCGCTACGATGAGAGTGCCAGACTCGGGCATC 2760
QY 2992 GACCTGCACTGACACTGCGCCCTGTATGCGAGCTTCGCTGAGCTGAGGCTCAAGCAT 3051
Db |||||
2761 GACCTGCACTGACACTGCGCCCTGTATGCGAGCTTCGCTGAGCTGAGGCTCAAGCAT 2820
QY 3052 GGCCAGCTGGAGAACAGGCCCTTAACCTCTCCACCGCCGGCTCCACACTGCCGAAA 3111
Db |||||
2821 GGCCAGCTGGAGAACAGGCCCTTAACCTCTCCACCGCCGGCTCCACACTGCCGAAA 2880
QY 3112 GCAGCCTTCCTGCTCGGTGCAAGATGCTGCTGCTGAAAACACAGGCTCAGCCTTCCAGG 3171
Db |||||
2881 GCAGCCTTCCTGCTCGGTGCAAGATGCTGCTGCTGAAAACACAGGCTCAGCCTTCCAGG 2940
QY 3172 GAATTCAGACCCCGGCTACAGTGGAAACAGGGCTCCAGCAGCAAGGTGGGCG 3231
Db |||||
2941 GAATTCAGACCCCGGCTACAGTGGAAACAGGGCTCCAGCAGCAAGGTGGGCG 3000
QY 3232 AAGCAGATGCTCCAGGATTTACACCTGAGCCCTGCCACCTGCTGAAAAAATCAT 3291
Db |||||
3001 AAGCAGATGCTCCAGGATTTACACCTGAGCCCTGCCACCTGCTGAAAAAATCAT 3060
QY 3292 CGGCCACGTGAAGAGACAGAAGGAGGATGGAGGATTAACCTGGGAAAACAAACAGGGA 3351
Db |||||
3061 CGGCCACGTGAAGAGACAGAAGGAGGATGGAGGATTAACCTGGGAAAACAAACAGGGA 3120
QY 3352 TCTTTTCTGCCCTGCTCCAGTCCAGTTGGCTGA 3387
Db |||||
3121 TCTTTTCTGCCCTGCTCCAGTCCAGTTGGCTGA 3156

RESULT 10
AAx87842
ID AAx87842 standard; DNA; 2844 BP.
XX
AC AAx87842;
XX
DT 09-NOV-1999 (first entry)
XX
DE NF-kB inducing kinase (NIK) DNA.
XX
KW NIK; NF-kB inducing kinase; inflammation; tumour necrosis factor;

KW interleukin-1; cytokine; inhibitor; antiinflammatory; apoptosis;
KW anti-apoptotic; human; ss.
XX
OS Homo sapiens.
XX
PN WO9943704-A1.
XX
PD 02-SEP-1999.
XX
PF 25-FEB-1999; 99WO-US004110.
XX
PR 27-FEB-1998; 98US-0076299P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Greene WC, Lin X, Gelezuinas R;
XX
DR WPI; 1999-518837/43.
XX
DR P-PSDB; AAY31665.
XX
PT New NF-kB inducing kinase proteins, used for inhibiting a kB-dependent
PT immune response, e.g. an inflammatory response or an anti-apoptotic
PT response.
XX
PS Disclosure; Fig 10; 48pp; English.
XX
CC This DNA sequence codes for human wild-type NIK (see AAY31665), a NF-kB
CC inducing kinase having serine/threonine kinase activity. The invention
CC relates to: (a) an N-terminus deletion mutant NIK protein (see AAY31666);
CC and (b) a kinase deficient NIK mutant protein (see AAY31667 and AAY31668)
CC that inhibits auto-phosphorylation or transphosphorylation. The invention
CC provides the molecular basis for cytokine induction of NF-kB-dependent
CC immune and inflammatory responses, emphasising a role for both NIK-NIK
CC and NIK-IKK (Ikb-specific kinase) interactions. A novel and highly
CC specific method for modulating NF-kB-dependent immune, inflammatory and
CC anti-apoptotic responses is based on interruption of the critical protein
CC interaction of NIK and IKK. The mutant NIK proteins are used in claimed
CC methods for inhibiting NF-kB-dependent gene expression. The kinase-
CC deficient NIK mutant proteins inhibit activation of IKK. The N-terminal
CC deletion NIK mutant proteins bind to IKK, thus inhibiting NIK/IKK
CC interaction
XX
SQ Sequence 2844 BP; 665 A; 887 C; 809 G; 483 T; 0 U; 0 Other;
Query Match 61.8%; Score 2839.2; DB 2; Length 2844;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2841; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 232 ATGGCAGTGTGGAATGGCTGCCAGGTGCCCTGCTCAGCAGTGGGGCAGCAGAAG 291
Db |||||
1 ATGGCAGTGTGGAATGGCTGCCAGGTGCCCTGCTCAGCAGTGGGGCAGCAGAAG 60
QY 292 GAATCCCCCAGCCAAAGAGAGACGCGCCACCTGGGGAGAAACAGAGCTCCGTCTAC 351
Db |||||
61 GAATCCCCCAGCCAAAGAGAGACGCGCCACCTGGGGAGAAACAGAGCTCCGTCTAC 120
QY 352 AGCTTTGAGGCGGTGGAGAGAGCCCTGTGTTCTCGGAAAGTGAGATCCTGATGAC 411
Db |||||
121 AAGCCGAGGCGGTGGAGAGAGCCCTGTGTTCTCGGAAAGTGAGATCCTGATGAC 180
QY 412 GTGATTACCAAGGGCAGCAGCCAAAGGAGCTCCGAGGAGGAGCTGCCATCTCTATC 471
Db |||||
181 GTGATTACCAAGGGCAGCAGCCAAAGGAGCTCCGAGGAGGAGCTGCCATCTCTATC 240
QY 472 ATCGCCAGGCTGAGTGTGAGAAATAGCCAAAGTTACGCCCCACCTTTTCAGAACGCATT 531
Db |||||
241 ATCGCCAGGCTGAGTGTGAGAAATAGCCAAAGTTACGCCCCACCTTTTCAGAACGCATT 300
QY 532 TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAT 591
Db |||||
301 TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAT 360
QY 592 GTGGCCCCATGCTACAGAGGGGCAAAATGGCCCGTGTGTGTTGGAAGGAAAGCGTCGACG 651

Db 361 GTGGCCCATGCTACAGAGGCAAAATGGCCGTGTGTGTGGAAGGGAAGAGCTGCGCAGC 420
Qy 652 AAAGCCCGAAGAAACGGAAGAAAGAGCTCAAAAGTCCCTGGCTCATGCGAGGAGTGCC 711
Db 421 AAAGCCCGAAGAAACGGAAGAAAGAGCTCAAAAGTCCCTGGCTCATGCGAGGAGTGCC 480
Qy 712 TTGGCCAAACCCCTCCCGAGACCCCTGAGCAGAGAGAGTGACCAATCCCAAGTGAGGAG 771
Db 481 TTGGCCAAACCCCTCCCGAGACCCCTGAGCAGAGAGAGTGACCAATCCCAAGTGAGGAG 540
Qy 772 GATGAGTCTCCACCTGGCGCCCAATATGTTAGAAACACCCCGAGTTCACCAAGCCTCTG 831
Db 541 GATGAGTCTCCACCTGGCGCCCAATATGTTAGAAACACCCCGAGTTCACCAAGCCTCTG 600
Qy 832 AAGGAACAGAGCCTTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGGGCTACGGCGGGCT 891
Db 601 AAGGAACAGAGCCTTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGGGCTACGGCGGGCT 660
Qy 892 CTGCTCGATCAGAACTCCAAACTGATCAGCCCTTTGCAATGTCGAAACCAAGTGAG 951
Db 661 CTGCTCGATCAGAACTCCAAACTGATCAGCCCTTTGCAATGTCGAAACCAAGTGAG 720
Qy 952 AAACTGACACACCCCGAGAGCGAGGCCCTGCCCCCTGCCCCAGCAGCCCTTCCCTAT 1011
Db 721 AAACTGACACACCCCGAGAGCGAGGCCCTGCCCCCTGCCCCAGCAGCCCTTCCCTAT 780
Qy 1012 AGCAGACTGCTCATCCCTTCCATTCCACCTCTCCAGCCCTGGAACCTCACCCTCTG 1071
Db 781 AGCAGACTGCTCATCCCTTCCATTCCACCTCTCCAGCCCTGGAACCTCACCCTCTG 840
Qy 1072 GAGTCTTCTCCCTGGGCAAACTGGCTGTGTAGACAGCCAGAAACCTTCCCTGAGCCACAC 1131
Db 841 GAGTCTTCTCCCTGGGCAAACTGGCTGTGTAGACAGCCAGAAACCTTCCCTGAGCCACAC 900
Qy 1132 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGAGAGCC 1191
Db 901 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGAGAGCC 960
Qy 1192 AGCTGCTGTCTCGTGGTGCCCATGAGAAATTTCTGTGGAGGAATACCTAGTGCATGCT 1251
Db 961 AGCTGCTGTCTCGTGGTGCCCATGAGAAATTTCTGTGGAGGAATACCTAGTGCATGCT 1020
Qy 1252 CTGCAAGGCAGCTGAGCTCAAGCCAGGCCACACAGCTGACAGCTTGGCCAAAGCCTGG 1311
Db 1021 CTGCAAGGCAGCTGAGCTCAAGCCAGGCCACACAGCTGACAGCTTGGCCAAAGCCTGG 1080
Qy 1312 GCAGCAGGGGCTCCAGATCCCGGAGCCCGAGCCCAAACTGAGGACAAAGAGGGTGTG 1371
Db 1081 GCAGCAGGGGCTCCAGATCCCGGAGCCCGAGCCCAAACTGAGGACAAAGAGGGTGTG 1140
Qy 1372 CTGCTCACTGAGAACTCAAGCCAGTGATATGAGTACCGAGAGAGTCCACTGGGCC 1431
Db 1141 CTGCTCACTGAGAACTCAAGCCAGTGATATGAGTACCGAGAGAGTCCACTGGGCC 1200
Qy 1432 ACGCACCAGCTCCGCTGGGAGAGGCTCTTTCGAGAGGTGCACAGATGAGAGACAAG 1491
Db 1201 ACGCACCAGCTCCGCTGGGAGAGGCTCTTTCGAGAGGTGCACAGATGAGAGACAAG 1260
Qy 1492 CAGACTGGCTTCCAGTGGCTGTCAAAAGAGTGGCCCTGGAAATTTTCGGGCGAGAGGAG 1551
Db 1261 CAGACTGGCTTCCAGTGGCTGTCAAAAGAGTGGCCCTGGAAATTTTCGGGCGAGAGGAG 1320
Qy 1552 CTGATGGCATGTGAGGATGAGCTCACCAAGAAATGTCCTTTGATGAGAGCTGTGAGA 1611
Db 1321 CTGATGGCATGTGAGGATGAGCTCACCAAGAAATGTCCTTTGATGAGAGCTGTGAGA 1380
Qy 1612 GAAGGGCTTGGGTCAAATCTTCAATGAGCTGTGGAAGTGTCTCCCTGGGCCAGCTG 1671
Db 1381 GAAGGGCTTGGGTCAAATCTTCAATGAGCTGTGGAAGTGTCTCCCTGGGCCAGCTG 1440
Qy 1672 GTCAAGGAGCAGGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTG 1731

Db 1441 GTCAGGAGCAGGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTG 1500
Qy 1732 GAGGGTCTGGAAATACCTCCACTCAGAAAGGATTTCTGCATGGGGAGCTCAAAAGCTGACAAC 1791
Db 1501 GAGGGTCTGGAAATACCTCCACTCAGAAAGGATTTCTGCATGGGGAGCTCAAAAGCTGACAAC 1560
Qy 1792 GTGCTCTCTGCTCCAGGATGGAGCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1851
Db 1561 GTGCTCTCTGCTCCAGGATGGAGCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1620
Qy 1852 CTTCAAACCTGATGCCCTGGGAAAGTCTTGTCTCAGGGGACTCATCCCTGGGCACAGAG 1911
Db 1621 CTTCAAACCTGATGCCCTGGGAAAGTCTTGTCTCAGGGGACTCATCCCTGGGCACAGAG 1680
Qy 1912 ACCCACATGGCTCCGAGAGTGGTGTGGGCAAGGAGCTGCGACGCAAGTGGATGCTGTGG 1971
Db 1681 ACCCACATGGCTCCGAGGTTGTGTGGGCAAGGAGCTGCGACGCAAGTGGATGCTGTGG 1740
Qy 1972 AGCAGTGTCTGTATGATGTGTCACATGCTCAAAGGCTGCCACCCCTGGACTCAGTTCTTC 2031
Db 1741 AGCAGTGTCTGTATGATGTGTCACATGCTCAAAGGCTGCCACCCCTGGACTCAGTTCTTC 1800
Qy 2032 CGAGGGCGCTCTGCTCAAGATTTGCCAGAGGCTCCGCTGTGAGGAGATCCACCC 2091
Db 1801 CGAGGGCGCTCTGCTCAAGATTTGCCAGAGGCTCCGCTGTGAGGAGATCCACCC 1860
Qy 2092 TCCTCGCGCCCTCTCACAAGCCAGGCCATCAAGAGGGGCTGAGGAAAGAGCCATCCAC 2151
Db 1861 TCCTCGCGCCCTCTCACAAGCCAGGCCATCAAGAGGGGCTGAGGAAAGAGCCATCCAC 1920
Qy 2152 CGCTGTCTGAGCGGAGCTGGGAGGAAGTGAACCGGGCACTACAGCAAGTGGAGGT 2211
Db 1921 CGCTGTCTGAGCGGAGCTGGGAGGAAGTGAACCGGGCACTACAGCAAGTGGAGGT 1980
Qy 2212 CTGAAGAGCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCC 2271
Db 1981 CTGAAGAGCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCC 2040
Qy 2272 AATTACCAACAGACCCCTCCATCCAGCCAGAGAGAGCTTTCCCAAGGGCCCGAGGGCC 2331
Db 2041 AATTACCAACAGACCCCTCCATCCAGCCAGAGAGAGCTTTCCCAAGGGCCCGAGGGCC 2100
Qy 2332 CGGCCAGCTGAGGAGACAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCTCCACACAGAG 2391
Db 2101 CGGCCAGCTGAGGAGACAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCTCCACACAGAG 2160
Qy 2392 CCCCAGAGCCAAACAAGTCTCTCTCTGATTTGAGCAAGGAGAGTCTGGGATGTGG 2451
Db 2161 CCCCAGAGCCAAACAAGTCTCTCTCTGATTTGAGCAAGGAGAGTCTGGGATGTGG 2220
Qy 2452 GAACCCCTTACTCTGTCTCTCCCTGAGCCAGCCCTGCGAGAAACCCAGCTCACACAGAG 2511
Db 2221 GAACCCCTTACTCTGTCTCTCTGAGCCAGCCCTGCGAGAAACCCAGCTCACACAGAG 2280
Qy 2512 CGGAAGCAACCCGTCGCGAGCAGGAATCTGAGCAGCTGGAAATAGAAATTTCTCTCAAC 2571
Db 2281 CGGAAGCAACCCGTCGCGAGCAGGAATCTGAGCAGCTGGAAATAGAAATTTCTCTCAAC 2340
Qy 2572 AGCTGTCTCCAGCCATTTCTCTGGAGAGCAGAGCAAAATTTCTCTGTGCTCAGCATC 2631
Db 2341 AGCTGTCTCCAGCCATTTCTCTGGAGAGCAGAGCAAAATTTCTCTGTGCTCAGCATC 2400
Qy 2632 GACAGCTCTCTCTGTGCGATGACAGTGAAGAAACCCATCAAAGGCTCTCAAGCTCG 2691
Db 2401 GACAGCTCTCTCTGTGCGATGACAGTGAAGAAACCCATCAAAGGCTCTCAAGCTCG 2460
Qy 2692 CGGACACCCCTGAGCTCAGGGGTACACTCTCTGAGCAGCCAGGGCGAGGCTCGAAGCTCC 2751
Db 2461 CGGACACCCCTGAGCTCAGGGGTACACTCTCTGAGCAGCCAGGGCGAGGCTCGAAGCTCC 2520
Qy 2752 AGCTGGAACATGTTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTT 2811
Db 2521 AGCTGGAACATGTTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTT 2580

QY 2812 GTGAAGTCAAATACAGTCTCTTAATGTTGAACACCTGCACATCCGGAGTTCCACCGG 2871
DB |||||
DB 2581 GTGAAGTCAAATACAGTCTCTTAATGTTGAACACCTGCACATCCGGAGTTCCACCGG 2640
QY 2872 GTCAAAGTGGGAGACATGCCACTGGCATCAGACGCCAGATCCCAGCTGCAGCTTCAGC 2931
DB |||||
DB 2641 GTCAAAGTGGGAGACATGCCACTGGCATCAGACGCCAGATCCCAGCTGCAGCTTCAGC 2700
QY 2932 TTGGTCAACCAAGACGGGAGCTGTTCCGTACGACATGGAGTCCAGACTCGGGCATC 2991
DB |||||
DB 2701 TTGGTCAACCAAGACGGGAGCTGTTCCGTACGACATGGAGTCCAGACTCGGGCATC 2760
QY 2992 GACCTGCAGTGACACTGGCCCCCTGATGGCAGCTTCGCCCTGGAGCTGGAGGGTCAAGCAT 3051
DB |||||
DB 2761 GACCTGCAGTGACACTGGCCCCCTGATGGCAGCTTCGCCCTGGAGTGGAGGGTCAAGCAT 2820
QY 3052 GCCCAGCTGGAGAAACAGGCCCTAA 3075
DB |||||
DB 2821 GCCCAGCTGGAGAAACAGGCCCTAA 2844

RESULT 11
ADQ67410
ID ADQ67410 standard; cdNA; 3152 BP.
AC ADQ67410;
XX
XX
DT 07-OCT-2004 (first entry)
XX
DE Novel human cdNA sequence #2383.
XX
KW ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX
XX Homo sapiens.
XX
XX EP1440981-A2.
XX
XX 28-JUL-2004.
XX
XX 21-JAN-2004; 2004EP-00001196.
XX
XX 21-JAN-2003; 2003JP-00102206.
XX
XX 09-MAY-2003; 2003JP-00131392.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX Yamamoto J, Isono Y, Nagai K, Irie R;
XX
XX WPI; 2004-535376/52.
XX
XX P-PSDB; ADQ67717.
XX
XX Novel 2495 cdNA, useful for treating osteoporosis, neurological diseases,
XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
XX Claim 1; SEQ ID NO 4571; 2449bp; English.
XX
XX The invention relates to 2495 novel polynucleotides (I) and their encoded
XX polypeptides, sequences hybridizing to these nucleotides, sequences
XX encoding partial polypeptides and sequences having 70% or 90% identity to
XX the nucleotide and protein sequences. The nucleotides and polypeptides
XX are useful as diagnostic markers or therapeutic target for the diseases
XX or morbid states. They are also useful for treating osteoporosis,
XX neurological diseases, Alzheimer's diseases, Parkinson's diseases,
XX dementia and various cancers. This sequence corresponds to a nucleotide
XX sequence of the invention.
XX
XX Sequence 3152 BP; 704 A; 1000 C; 874 G; 574 T; 0 U; 0 Other;

Query Match 58.2%; Score 2673.4; DB 12; Length 3152;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 36; Indels 10; Gaps 7;
QY 1764 TCTGATGGGAGACGTCAAAGCTGACAACGTGCTCTGTCCAGCGATGGAGCCAGCGAC 1823
DB |||||
DB 340 TCCCGGGGGCGGCTGGCAGCTGACAACGTGCTCTGTCCAGCGATGGAGCCAGCGAC 399
QY 1824 CCTCTGTGACTTTGGCCATGCTGTGCTTCAACTGATGGCCCTGGGAAAGTCTCTTGT 1883
DB |||||
DB 400 CCTCTGTGACTTTGGCCATGCTGTGCTTCAACTGATGGCCCTGGGAAAGTCTCTTGT 459
QY 1884 CACAGGGGACTACATCCTTGGCAGACAGACCCACATGGCTCCGAGGTGGTCTGGGAG 1943
DB |||||
DB 460 CACAGGGGACTACATCCTTGGCAGACAGACCCACATGGCTCCGAGGTGGTCTGGGAG 519
QY 1944 GAGCTGCAGCGCCAAAGTGGATGTCTGGAGCAGCTGTGTATGATGTGCAATGCTCAA 2003
DB |||||
DB 520 GAGCTGCAGCGCCAAAGTGGATGTCTGGAGCAGCTGTGTATGATGTGCAATGCTCAA 579
QY 2004 CGGCTGCCACCCCTGGACTCAGTTCTCCGAGGGCGCTCTGCCCTCAAGATTGCCAGCA 2063
DB |||||
DB 580 CGGCTGCCACCCCTGGACTCAGTTCTCCGAGGGCGCTCTGCCCTCAAGATTGCCAGCA 639
QY 2064 GCCTCCGCTGTGAGGGAGATCCCACTCTCGGCCCTCTTCACAGCCAGGCCATCCA 2123
DB |||||
DB 640 GCCTCCGCTGTGAGGGAGATCCCACTCTCGGCCCTCTTCACAGCCAGGCCATCCA 699
QY 2124 AGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGGAGGGAGCTGGGAGGAAGT 2183
DB |||||
DB 700 AGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGACGGGAGCTGGGAGGAAGT 759
QY 2184 GAACCGGGCACTACAGCAAGTGGGAGTCTGAAGAGCCCTTGGAGGGAGAAATAAAGA 2243
DB |||||
DB 760 GAACCGGGCACTACAGCAAGTGGGAGTCTGAAGAGCCCTTGGAGGGAGAAATAAAGA 819
QY 2244 ACCAAGACATCCACCGCCAAATCAAGCCAAATTAACACAGAGCCCTTCCATGCCAGCCGAG 2303
DB |||||
DB 820 ACCAAGACATCCACCGCCAAATCAAGCCAAATTAACACAGAGCCCTTCCATGCCAGCCGAG 879
QY 2304 AGAGCTTTGCGCAAGGGCCCCAGGGCCCCGGCCAGCTGAGGAGACAACAGGAGAGCCCC 2363
DB |||||
DB 880 AGAGCTTTGCGCAAGGGCCCCAGGGCCCCGGCCAGCTGAGGAGACAACAGGAGAGCCCC 939
QY 2364 TAAAGCTCCAGCTCTCTCCACAGAGCCCCCAGAGCAACAAGTCTCTCTCCCTTGAC 2423
DB |||||
DB 940 TAAAGCTCCAGCTCTCTCTCCACAGAGCCCCCAGAGCAACAAGTCTCTCTCCCTTGAC 999
QY 2424 TTTGAGCAAGGAGGAGTCTGGGATGTGGAAACCTTACCTCTGTCTCTCCCTGGAGCCAGC 2483
DB |||||
DB 1000 TTTGAGCAAGGAGGAGTCTGGGATGTGGAAACCTTACCTCTGTCTCTCCCTGGAGCCAGC 1059
QY 2484 CCTTCCAGAAAACCCAGCTACAGAGGGGAAAGAACCCGTCCCGGAGCAGAACTGCA 2543
DB |||||
DB 1060 CCTTCCAGAAAACCCAGCTACAGAGGGGAAAGAACCCGTCCCGGAGCAGAACTGCA 1119
QY 2544 GCAGCTGGAAATAGAAATTTCTCAACAGCCCTGTCCAGCCATTTTCTCTGAGGAGCA 2603
DB |||||
DB 1120 GCAGCTGGAAATAGAAATTTCTCAACAGCCCTGTCCAGCCATTTTCTCTGAGGAGCA 1179
QY 2604 GGAGCAAAATTTCTCTCGTGCCTCAGCATCGACAGCCCTCTCCCTGTCCGATGACAGTGA 2663
DB |||||
DB 1180 GGAGCAAAATTTCTCTCGTGCCTCAGCATCGACAGCCCTCTCCCTGTCCGATGACAGTGA 1239
QY 2664 GAACCCATCAAAGGCTCTCAAAGCTCGGGGACACCTTGAGCTCAGGGGTACACTCTG 2723
DB |||||
DB 1240 GAACCCATCAAAGGCTCTCAAAGCTCGGGGACACCTTGAGCTCAGGGGTACACTCTG 1299
QY 2724 GAGCAGCCAGGCCGAGGGCTCGAAGCTCCAGCTGGAAACATGGTGTGGCCCCGGGGCGCC 2783
DB |||||
DB 1300 GAGCAGCCAGGCCGAGGGCTCGAAGCTCCAGCTGGAAACATGGTGTGGCCCCGGGGCGCC 1359
QY 2784 CACCGACACCCCAAGCTATTTCAATGGTGTGAAGTCCAAATACAGTCTCTTTAATGGTGA 2843

```
Db 1360 CACCCACACCCCAAGCTATTTCAATGGTGTGAAAGTCCAAATACAGTCTCTTAATGGTGA 1419
Qy 2844 ACACCTGCACATCCCGGAGTCCACCGGGTCAAAAGTGGGAGACATCGGCCACTGGGCATCAG 2903
Db 1420 ACACCTGCACATCCCGGAGTTCACCGGGTCAAAAGTGGGAGACATCGGCCACTGGGCATCAG 1479
Qy 2904 CAGCCAGATCCAGCTGAGCTTCAGCTTGGTACAAAGACGGGAGCCTGTTGGCTA 2963
Db 1480 CAGCCAGATCCAGCTGAGCTTCAGCTTGGTACAAAGACGGGAGCCTGTTGGCTA 1539
Qy 2964 CGACATGAGGTGCCAGACTCGGCATCGACTCGAGTGCACCTGCGAGTGCACCTGCGCCCTGATGGCAG 3023
Db 1540 CGACATGAGGTGCCAGACTCGGCATCGACTCGAGTGCACCTGCGAGTGCACCTGCGCCCTGATGGCAG 1599
Qy 3024 CTTGCGCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTTAACCCCTGCCC 3083
Db 1600 CTTGCGCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTTAACCCCTGCCC 1659
Qy 3084 TCCACCGCGGCTCCACACTGCGCGAAAGAGCCTTCTGCTCGGTGCAACATGCTGCC 3143
Db 1660 TCCACCGCGGCTCCACACTGCGCG--AAGCAGCCTTCTGCTCGGTGCAACATGCTGCC 1718
Qy 3144 TGAACAACAGGCTCAGCCTTCCAGGGGATTGCCAGCCCGCCCGCTCAGTGGGAAC 3203
Db 1719 TG-AAACAACAGGCTCAGCCTTCCAGGGATTGCCAG--CCCCCGGCTCAGCAGTGGGAC 1776
Qy 3204 CAGGCGCTCGCAGCAGCAGGTGGGGCAAGCAGAAATGCTCCAGGATTTCAACCTGA 3263
Db 1777 CAGGCGCTCGCAGCAGCAGGTGGGGCAAGCAGAAATGCTCCAGGATTTCAACCTGA 1836
Qy 3264 GCCCTGCCCGCTGCTGCTGAAACAA-TCGCCACAGTGAAGAGACAGAGGAGATGGC 3322
Db 1837 GCCCTGCCCGCTGCTGCTGAGAAACACTCCGCCACAGTGAAGAGACAG-AGGAGGATGGC 1895
Qy 3323 AGGAGTTACCTGGGGAAACAAACAGGGATCTTTTCTGCCCTGCTCAGTCCAGTTGG 3382
Db 1896 AGGAGTTACCTGGGGAAACAAACAGG---ATCTTCTGCCCTGCTCAGTCCAGTTGG 1951
Qy 3383 CCTGACCCGCTTGGATCAGTCAACATTTGTTGGCAGACAGGGAGCAGCTTCCAGCCT 3442
Db 1952 CCTGACCCGCTTGGATCAGTCAACATTTGTTGGCAGACAGGGAGCAGCTTCCAGCCT 2011
Qy 3443 GGGTCAGAAAGGGTGGGCGAGCCCTTCGGCCCGCTCACCCTCCAGGCTGCTGTGAGAGTGT 3502
Db 2012 GGGTCAGAAAGGGTGGGCGAGCCCTTCGGCCCGCTCACCCTCCAGGCTGCTGTGAGAGTGT 2071
Qy 3503 CAAGTGTGTAAGGGCCCAAACTCAGGTTCAAGTGCAGAACAGGTCAGAGGTATGCCCGC 3562
Db 2072 CAAGTGTGTAAGGGCCCAAACTCAGGTTCAAGTGCAGAACAGGTCAGAGGTATGCCCGC 2131
Qy 3563 CCGTAGGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCTCAGCTGGCCAGCTCACCCCT 3622
Db 2132 CCGTAGGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCTCAGCTGGCCAGCTCACCCCT 2191
Qy 3623 TTTGGGTGTAGGGGAAAAAGATGCTGACCTCGGGAAGGCTCCCTGGTAGAATACACCAC 3682
Db 2192 TTTGGGTGTAGGGGAAAAAGATGCTGACCTCGGGAAGGCTCCCTGGTAGAATACACCAC 2251
Qy 3683 ACTTTTCAAGTTGTGAAACACAGGTCCTGAGTTGACCTCTGGTTCCAGCCAAAGACCAAA 3742
Db 2252 ACTTTTCAAGTTGTGAAACACAGGTCCTGAGTTGACCTCTGGTTCCAGCCAAAGACCAAA 2311
Qy 3743 GAAGGTGTGAAGTGAAGTGTCTCAGTCCCAACATGTGCCCTTTGCTGCTGGCTA 3802
Db 2312 GAAGGTGTGAAGTGAAGTGTCTCAGTCCCAACATGTGCCCTTTGCTGCTGGCTA 2371
Qy 3803 CCACCTTTCCCGCAGACAGCAGGCCCGAGCCCTTCCAGGCCACAGCTGCCCGCAGACTC 3862
Db 2372 CCACCTTTCCCGCAGACAGCAGGCCCGAGCCCTTCCAGGCCACAGCTGCCCGCAGACTC 2431
Qy 3863 GCTGCGACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTGATGCGGATATGCTTGCACAGG 3922
```

```
Db 2432 GCTGGCACTCAGTTCCCTCATCTGCTAAAGGTGAAGGGTGATGCAAGATATGCTTGACAGG 2491
Qy 3923 AACAGTCTGTGGATGGACATGATCAGTCTAAGAAAGACAGCAGAGAGAGACG-TCGCGC 3981
Db 2492 AACAGTCTGTGGATGGACATGATCAGTCTAAGAAAGACAGCAGAGAGAGACGCTCCGCGC 2551
Qy 3982 GCCCCAGCCCCACTATCAGTGTCCAGGTGCTGCTTCCCCAGAGACAGCTCAGCATCAC 4041
Db 2552 GCCCCAGCCCCACTATCAGTGTCCAGGTGCTGCTTCCCCAGAGACAGCTCAGCATCAC 2611
Qy 4042 ACTCAGCACTACCCCTGCCCCTGGCCAGAGGGTACTTCCGACGGCACTTTGCACTC 4101
Db 2612 ACTCAGCACTACCCCTGCCCCTGGCCAGAGGGTACTTCCGACGGCACTTTGCACTC 2671
Qy 4102 TGATGACCTCAAAGCACTTTTCATGGCTCTGGCAGGGCAGGGCAGGGCAGTGAACAC 4161
Db 2672 TGATGACCTCAAAGCACTTTTCATGGCTCTGGCAGGGCAGGGCAGGGCAGTGAACAC 2731
Qy 4162 TGTAGGAGCATAGCAAGCCAGGAGATGGGTGAAGGGACACAGTCTTTGAGCTGTCCACAT 4221
Db 2732 TGTAGGAGCATAGCAAGCCAGGAGATGGGTGAAGGGACACAGTCTTTGAGCTGTCCACAT 2791
Qy 4222 GCATGTGACTCTCAAACCTCTTCAGATTTCTTAAGAAATAGCACCCCTTCCCCATTG 4281
Db 2792 GCATGTGACTCTCAAACCTCTTCAGATTTCTTAAGAAATAGCACCCCTTCCCCATTG 2851
Qy 4282 CCCAGCTTACGCTCTTTCTCCAGGGAGCTTCTCAGGACTCACGTAGCATTAATACAGC 4341
Db 2852 CCCAGCTTACGCTCTTTCTCCAGGGAGCTTCTCAGGACTCACGTAGCATTAATACAGC 2911
Qy 4342 TGTGAATCGTCAGGGGTGCTGTAGCTCAACCTCTCTGGGGCAGGGGACGCCGAGACT 4401
Db 2912 TGTGAATCGTCAGGGGTGCTGTAGCTCAACCTCTCTGGGGCAGGGGACGCCGAGACT 2971
Qy 4402 CCGTGGGAGAGCTCATTTCCACATCTTGCCAGACAGCCTTTGTCAGCTGTCCACATT 4461
Db 2972 CCGTGGGAGAGCTCATTTCCACATCTTGCCAGACAGCCTTTGTCAGCTGTCCACATT 3031
Qy 4462 GAGTCAGACTGCTCCCGGGAGAGAGCCCGGGCCCCCAGCACAATAAGAACTGACAGCCTT 4521
Db 3032 GAGTCAGACTGCTCCCGGGAGAGAGCCCGGGCCCCCAGCACAATAAGAACTGACAGCCTT 3091
Qy 4522 GGTACTCGCAGAGTCTGGGTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGATGA 4581
Db 3092 GGTACTCGCAGAGTCTGGGTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGATGA 3151
Qy 4582 C 4582
Db 3152 C 3152

RESULT 12
ACN44437
ID ACN44437 standard; cDNA; 2828 BP.
XX
XX ACN44437;
XX AC
XX 18-NOV-2004 (first entry)
XX
XX Mouse mRNA sequence mC171748.
DE
XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ds.
XX
XX Mus musculus.
XX
XX WO2003073826-A2.
XX
XX 12-SEP-2003.
XX
XX 28-FEB-2003; 2003WO-US006235.
XX
XX 01-MAR-2002; 2002US-00087192.
XX
```

PA (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-328604/31.

DR Recombinant nucleic acid useful for diagnosis and treatment of carcinoma

XX comprises a nucleotide sequence.

PT Claim 1; SEQ ID NO 884; Opp; English.

XX The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA gene
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published

XX SQ Sequence 2828 BP; 688 A; 880 C; 790 G; 470 T; 0 U; 0 Other;

Query Match 44.4%; Score 2042.6; DB 11; Length 2828;

Best Local Similarity 83.4%; Pred. No. 0;

Matches 2377; Conservative 0; Mismatches 444; Indels 28; Gaps 4;

QY	232	ATGGCAGTGCATGAAATGCCCTGCCAGGTGCCCTTGGCTCAGCAGTGGGGCAGCAGAAG	291
DB	1	ATGGCCGTGATGGAAGTGCCTGCCCGGCATCTCTGGTACAGCTCGGGCAGCAGAAG	60
QY	292	GAACTCCCAAGCAAGAGAGAGAGCGCCCACTGGGGAAGAAACAGAGCTCCGTCTAC	351
DB	61	GAGCTTGCCAAAGCCAAAGAGAGACACAGTCACTGGGGAAGAGCAGAGTGCATCTTC	120
QY	352	AAGCTTGAGCCGTGGAGAGAGCCCTGTGTTCTGGGAAAGTGGAGATCTGATGAC	411
DB	121	AAGCTTGAGCCGTGGAGAGAGCCCTGTGTTCTGGGAAAGTGGAGATCTGATGAC	180
QY	412	GTGATTACCAAGGACACCCAGGAAGGCTCCGAGGAGGCGCAGCTCCATCTCTATC	471
DB	181	GTGATCAACCAAGGACACCCAGGAAGGCTCTGAGGAGGAGACACCGGCATCTCCATC	240
QY	472	ATCGCCCAAGGCTGAGTGTGAGAAATAGCCAAAGATTGAGTTCAGAACGCAAT	531
DB	241	ATCGCCCAAGGCTGAAATGTGAGAAATAGCCAAAGATTGAGTTCAGAACGCAAT	300
QY	532	TTCATCGTGGGTCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCAACAT	591
DB	301	TTCATCGGCGGTTCACAGCAGTACAGCCAGTCTGAGAGTCTCGATCAAAATCCCAACAT	360
QY	592	GTGGCCCATGCTACAGAGGCAAAATGGCCCTGTGTTGGAAGGAAGGCTCGCAGC	651
DB	361	GTGGCCCATGCACTGAGGCAAAATGGCCCTGTGTTGCGGAGGGAAGAGTCAAGGC	420
QY	652	AAAGCCCGGAAGAAACGGAAGAGAGAGTCAAAAGTCCCTGGCTCATGAGAGTGGCC	711
DB	421	AAAGCCCGGAAGAAACGTAAGGAAGAGAGTCAAAAGTCCCTGGCTCATGAGAGTGGCC	480
QY	712	TTGGCCCAACCCCTCCCAAGGACCCCTGAGCAGGAGAGTGCACCATCCAGTGCAGAG	771
DB	481	TTAGCCCAAGCCCTCCCAAGGACCCCTGAGCAGGAGAGTGCACCATCCAGTGCAGAG	540
QY	772	GATGAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCTCTG	831
DB	541	GATGAGTCTCCACTAGGCAACCTCTATGCCAGAAATGCTCCAGGTTTACCAAGCTCTG	600
QY	832	AAGGAACAGGCTTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGGCTACGGCCGGT	891

DB	601	GGGGACAGGCTTGGCCACCTGTGCTTTAAGAAACAAGATGAAGGCTGCGACGGTA	660
QY	892	CTGCTCGATCAGAACTCCAAACTGATCAGCCCTTGCATGTCATGAAACACGCTGG	951
DB	661	CTGCTCGACAGAACTCCAAACTGATCAGCCCTTGCATGTCATGAAACACGCTGG	720
QY	952	AAACTGCAACACCCCGAGAGGAGGCGCCCTGCGCCCTGCCACGACCCCTTCCCTAT	1011
DB	721	AAACTGCAACACCCCGAGGAGGCGCCCTGCGCCCTGCCACGACCCCTTCCCTAT	780
QY	1012	AGCAGCTGCTCATCTCCCTTCCATTCACCTCTCCAGCCCTGGAACCTCACCTCTG	1071
DB	781	AGCGAATGCCCCATCTCTTCCATTCACCTCTGGAGCCCTGGAACCTTATATGCTG	840
QY	1072	GAGT---CCTTCTGGGCAAACTGCGCTGTGTAGACAGCAGAAACCTTGGCTGACCA	1128
DB	841	GACTCTGCGCTCTGGCAAACTAGCGGCTGTACGCGCCAGCGGCTCTGCTGGCCCA	900
QY	1129	--CACCTGAGCAAACTGCGCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTG	1185
DB	901	CGCATCTAAGCAACTGCGCCCATGGAGACAGTCAAGAGCCGCTGCTGGCCACACCTG	960
QY	1186	GAGCCAGCTGCTCTCTGCTGTGCCCATGAGAAATTTCTGTGGAGGAATACCTAGTG	1245
DB	961	GAGTCCAGCTGCGCGTCTCGGGTGCCCTAGAAAGGTTCCCGTGGAGGAATACCTG	1020
QY	1246	CATGCTCTGCAAGCAGCGTGTAGCTCAAGCCAGGCCCCACAGCTGACAGCTGCCCCA	1305
DB	1021	CATGCGCTCCAAAGGAAGTGTGAGCTCAGGCCAGGCCCAAGCTGCGCAGCTGCTA	1080
QY	1306	ACCTGGGAGCAGCGGGCTCCAGATCCCGGAGGCCAGCCCCCAAACTGAGGACAAAC	1365
DB	1081	ACATGCTCTCGGAGAGGCCAGCTGAGGCTCGGCCCGGAACTGAGGACAAACGAG	1140
QY	1366	GCTGCTCTCTCACTAGAAAACCTCAAGCCAGTGGATATGAGTACCGAGAAAGTCCAC	1425
DB	1141	GGGCTGCTCTTACTAGAAAACCTCAAGCCAGTGGATATGAGTACGAGAAAGAGTCC	1200
QY	1426	TGGGCGCAGCAGCTCGGCTGGGAGAGGCTCTTGGAGAGGTCACAGAGTGGAG	1485
DB	1201	TGGATGACACACACAGCCCTCGGGTGGGAGAG--GCTCTTTCGGCGAGGTCACAG	1259
QY	1486	GACAGCAGACTGGCTTCCAGTGGCTGCTCAAAAGAGTGGCGCTGGAAGTATTTCCG	1545
DB	1260	GACAGCAGACAGGCTTCCAGTGGCTGCTCAAAAGAGTACGACTCGAGTGTTCGG	1319
QY	1546	GAGGAGCTGATGGCATGTGAGGATGACCTCAACAGAAATGTCCTTTGTATGAGCT	1605
DB	1320	GAGGAATAGTGGCTGTGCTGAGCTCGGCCAGAAATCGTCCCTCTCTATGAGCT	1379
QY	1606	GTGAGAGAGGCGCTTGGGTCAACATCTTCATGAGCTGCTGGAAGTGGCTCCCTGG	1665
DB	1380	GTGAGAGAGGCGCTTGGGTGAACATCTTCATGGAACCTGCTAGAAAGTGGCTCG	1439
QY	1666	CAGCTGCTGCAAGGAGCAGGCTGCTCCAGAGGAGCGGGCCCTGCTACTCTGCGG	1725
DB	1440	CAGCTCAAAAGCAAAATGGGTGTCTGCAAGAGCCAGCCCTTACTACTCTGGG	1499
QY	1726	GCTTGGAGGCTTGGAAATCCTCACTCAAGAAAGATTTGTCATGGGAGCTCAAGCT	1785
DB	1500	GCCCTGGAGGCGCTGGAGTACCTCCACACACGAGGATTTCTGATGGCGATGTCAA	1559
QY	1786	GACACGCTCTGCTTCCAGGATGGGAGCCACGAGCCCTCTGTGATTTGGCCATGCT	1845
DB	1560	GACAACTGCTCTGCTTCCAGTGTGGAAGCCGAGCGCCCTCTGCGACTTTGGCC	1619
QY	1846	GTGCTCTTCAACCTGATGGCTGGGAAAGTCTTGTCTCACAGGGAAGTACATCTCG	1905
DB	1620	TTGTGCTTCCAACTGAGCGCTAGGGAATCTTGTCTCACAGGGAAGTACATCTCG	1679
QY	1906	ACAGAGACCCACATGGCTTCCGAGGTGTGTGGGAGGAGCTGCGACGCGCAAGTGA	1965

Db 1580 ACGAGACCCACATGGCACAGAAAGTGGTGTATGGAAAGCCCTCGATGCCAAGTGGAC 1739
Qy 1566 GTCTGGAGCAGCTCTGTATGATGCTGCACATGCTCAACGGCTCCACCCCTGGACTCAG 2025
Db 1740 ATCTGGAGCAGCTCTGTATGATGCTGCACATGCTCAACGGCTCCACCCCTGGACTCAG 1759
Qy 2026 TTCTTCCGAGGGCCGCTCTGCTCAAGATTGCCAGCGAGCTCCGCTGTGAGGGAGATC 2085
Db 1800 TACTTCCGAGGGCCGCTTTGTCTCAAGATTGCCAGCGAGCTCCACCGATCAGGGAGATC 1859
Qy 2086 CCACCTCTCTCGGCCCTCTCACAGCCAGCGCCATCAAGAGGGCTGAGGAAAGAGGCC 2145
Db 1860 CCACCTCTCTCGGCCCTCTCACAGCCAGCGCCATCAAGAGGGCTGAGGAAAGAGGCC 1919
Qy 2146 ATCCACCGCTGTCTGACGCGAGCTGGAGGGAGGTGAACCGGGCACTACAGCAAGTG 2205
Db 1920 GTCCACCGAGCATCTGCCATGGAGCTTCGGAGGAAAGTGGCAAGGCACCTACAGGAAGTG 1979
Qy 2206 GGAGGTCTGAAGAGCCCTTGGAGGGAGATATATAAGAACCAAGACATCCACCGCCAAAT 2265
Db 1980 GGAGGTCTGAAGAGCCCTTGGAGGGAGATATATAAGAACCAAGACCTCCACCGCCAAAG 2039
Qy 2266 CAAGCCAAATTACCAACAGACCTCTCATGCCAGCCGAGAGAGCTTCGCCAAGGGCCCCA 2325
Db 2040 CAAGCCACTGCCACAGACCTTACTCTCCGCGAGAGA----- 2080
Qy 2326 GGGCCCGCGCAGCTGAGGAGACAAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCTCCCA 2385
Db 2081 --GAACCCACAGCCAGGCCCAACAGAGCGGGCTCTGAGCCTCAGCCTCTCTACCG 2138
Qy 2386 CCAGAGCCCCAGAGCCCAAGATCTCTCTCTGAGCTTTGAGCAAGAGGAGTCTGGG 2445
Db 2139 CCAGAACCCAGAACCCGAGCAAGAGCGCCAGCCCTGAACCTGAGCAAGGAGGAGTCTGGC 2198
Qy 2446 ATGTGGAAACCTTACTCTCTCTCTCTGAGCGAGCCCTGCGAGAACCCAGCTCA 2505
Db 2199 ACATGGAAACCCCTGCTCTCTCTCTGAGCCAGCCACTGCGCAAGGCCCGAGCTTC 2258
Qy 2506 CCAGAGGGGAAGCAACCGTCCCGGAGCAGGAATGCGAGAGCTGGAAATAGATTATTC 2565
Db 2259 CCAGACCGGAGGCAACCTTGCCAGAGCTGGAGCTACAGCAACTGGAGATAGACTGTTT 2318
Qy 2566 CTCAACAGCTGTCCAGCCATTTCTCTGAGGAGCAGGAGCAAAATCTCTGTGCTC 2625
Db 2319 CTCAACAGCTGTCCAGCCGTTCTCTGAGGAAACAGGAACAAATCTCTCTGTGCTC 2378
Qy 2626 AGCATGACAGCTCTCTCTCTGAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 2685
Db 2379 AGCATGACAGCTCTCTCTCTGAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 2438
Qy 2686 AGCTCGGGGACACCTGAGCTCAGGGCTACACTCTGAGCAGCGCGAGGCTCGA 2745
Db 2439 AGCTCAGGGACACCTGAGTCTGGGTGACCTTTGGAAACACAGCTGAGGCAAGA 2498
Qy 2746 AGCTCAGCTGGAATAGTGTCTGGCCGGGGGGGGCCACCGACACCCCAAGTATTTTC 2805
Db 2499 ACCTGAGCTGACAGCGGCTGGCCGGGGGGGGCTACTGACATCCCGAGCTACTTC 2558
Qy 2806 AATGGTGAAGTCCAAATACAGTCTTAAATGGTGAACCTGCAATCCGGAGGATTC 2865
Db 2559 AACGGGTCAAGGTCCAGATCAGTCTCTCAATGGCGAAACACCTTCATATCCGGGAATTC 2618
Qy 2866 CACCGGTCAAGTGGGAGACATCGCACTGGCATCAGCAGCCAGATCCAGCTGCAGCC 2925
Db 2619 CACCGGTCAAGTGGGAGACATTCGCAACCGGCATCAGCAGCCAGATCCAGCCACAGCT 2678
Qy 2926 TTCAGTGTGTTCACCAAGAGCGGAGCCTGTTCTGCTACGACATGAGAGGTGCCAGACTCG 2985
Db 2679 TTCAGCTGTGTGACCAAGATGGACAGCCTGTTTGTATGATGGAGGTGCCAGACTCG 2738
Qy 2986 GGCATCGACCTGCAAGTGCACATCGGCCCTCTGATGGCAGCTTCGCTGAGCTGAGGGGTC 3045
Db 2739 GGCATCGACCTGCAAGTGCACCTCGGCCCTCTGATGGCAGCTTTGTGTCAGCTGAGGGTC 2798

Qy 3046 AAGCATGGCAGCTGGAGAAACAGGCCCTA 3074
Db 2799 AAGCATGGTCACTGGAGAACCGACCTA 2827

RESULT 13

AAV03324
ID AAV03324 standard; cDNA; 2631 BP.

XX AAV03324;

XX AC AAV03324;

XX DT 15-APR-1998 (first entry)

XX XX Clone 10 cDNA encoding NMPI, a TRAF2 binding protein.
XX Human tumour necrosis factor receptor-associated factor 2; TRAF2;
XX TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;
XX intracellular signalling activity; acute hepatitis;
XX autoimmune-induced cell death; ss.

XX OS Homo sapiens.

XX PN W09737016-A1.

XX PD 09-OCT-1997.

XX PF 01-APR-1997; 97WO-IL000117.

XX PR 02-APR-1996; 96IL-00117800.

XX PR 26-AUG-1996; 96IL-00119133.

XX PA (VEDA) YEDA RES & DEV CO LTD.

XX PI Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;

XX DR WPI; 1997-503101/46.

XX PT DNA encoding tumour necrosis factor receptor-associated factor binding
XX molecule - used for modulation or mediation in cells of the activity of
XX NF-kB.

XX PS Claim 4; Fig 4; 127pp; English.

XX CC The present sequence encodes a protein designated NMPI, a TRAF2 binding
XX protein. This protein contains Ser/Thr protein kinase motifs. A full
XX length cDNA clone encoding NF-kappaB (NIK) was obtained by PCR using the
XX present clone, clone 10. The clone 10 protein is capable of binding to at
XX least amino acids 222-501 of TRAF2. The TRAF-2 binding proteins can be
XX used for modulation or mediation in cells of the activity of NF-kappaB or
XX any other intracellular signalling activity modulated or mediated by
XX TRAF2. TRAF-binding proteins are especially used for prevention or
XX treatment of pathological conditions associated with NF-kB induction,
XX e.g. acute hepatitis, autoimmune-induced cell death, e.g. death of the
XX beta langerhans cells or the pancreas that results in diabetes, the death
XX of cells in graft rejection, the death of oligodendrocytes in the brain
XX in multiple sclerosis, and AIDS-inhibited T cell suicide which causes
XX proliferation of the AIDS virus and hence the AIDS disease. The proteins
XX are also useful for screening of ligands capable of binding to a protein,
XX which are useful for modulating cellular activity modulated/mediated by
XX TRAF2

SQ Sequence 2631 BP; 596 A; 791 C; 704 G; 459 T; 0 U; 81 Other;

Query Match 40.3%; Score 1851.2; DB 2; Length 2631;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2335; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

Qy 2099 CCCCCTCTCACAGGCCAGGCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 2158

Db 1 CCCCCTCTCACAGGCCAGGCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60

Qy 2159 CTGACGCGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 2218

Db 61 |||||CTGCACGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120
QY 2219 GCCCTTGGAGGGAGGAATATAAGAACCAAGACATCCAACGCAAAATCAAGCAATACC 2278
Db 121 GCCCTTGGAGGGAGGAATATAAGAACCAAGACATCCAACGCAAAATCAAGCAATACC 180
QY 2279 ACCAGACCTTCCATGCCACGACGAGAGCTTTGGCCAAAGGGCCCCCAGGGCCCCGGCCAG 2338
Db 181 ACCAGACCTTCCATGCCACGACGAGAGCTTTGGCCAAAGGGCCCCCAGGGCCCCGGCCAG 240
QY 2339 CTGAGGAGACAAACAGGACAGAGCCCTTAAGCTCCAGCCTCTCTCCACACAGAGCCCCCAG 2398
Db 241 CTGAGGAGACAAACAGGACAGAGCCCTTAAGCTCCAGCCTCTCTCCACACAGAGCCCCCAG 300
QY 2399 AGCCAAACAAAGTCTCTCCCTTGAATTTGAGCAAGGAGAGTCTGGGATGTGGGAACCTT 2458
Db 301 AGCCAAACAAAGTCTCTCCCTTGAATTTGAGCAAGGAGAGTCTGGGATGTGGGAACCTT 360
QY 2459 TACCTCTGTCTCCCTGGAGCAGACCCCTGCGCAGAAACCCAGCTCACAGAGCGGAAG 2518
Db 361 TACCTCTGTCTCCCTGGAGCAGACCCCTGCGCAGAAACCCAGCTCACAGAGCGGAAG 420
QY 2519 CAACCGTCCCGGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATATTTCTCAACAGCCTGT 2578
Db 421 CAACCGTCCCGGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATATTTCTCAACAGCCTGT 480
QY 2579 CCGAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATCGACAGCC 2638
Db 481 CCGAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATCGACAGCC 540
QY 2639 TCTCCCTGTGCGATGACAGTAGAGAACCCATCAAGGCCCTCAAGCTCGCGGAC 2698
Db 541 TCTCCCTGTGCGATGACAGTAGAGAACCCATCAAGGCCCTCAAGCTCGCGGAC 600
QY 2699 CCTGAGCTCAGGCGTACACTCTCTGAGCAGCAGCGCCGAGGCTCGAAGCTCCAGCTGA 2758
Db 601 CCTGAGCTCAGGCGTACACTCTCTGAGCAGCAGCGCCGAGGCTCGAAGCTCCAGCTGA 660
QY 2759 ACATGCTGTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAG 2818
Db 661 ACATGCTGTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAG 720
QY 2819 TCCAAATACAGTCTCTTAATGTGTGAACCTTCAATCGGAGTTCCACCGGGTCAAAG 2878
Db 721 TCCAAATACAGTCTCTTAATGTGTGAACCTTCAATCGGAGTTCCACCGGGTCAAAG 780
QY 2879 TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTGCAGCCTTCAGCTTGGTCA 2938
Db 781 TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTGCAGCCTTCAGCTTGGTCA 840
QY 2939 CCAAGACGGGAGCCTGTTCCGTAGCATGGAGGTGCGACATCGGGCATCGACCTGC 2998
Db 841 CCAAGACGGGAGCCTGTTCCGTAGCATGGAGGTGCGACATCGGGCATCGACCTGC 900
QY 2999 AGTGACACTGCGCCCTGATGCGAGCTTCGCTTGGAGCTGGAAGGTCAAGCATGSCCAGC 3058
Db 901 AGTGACACTGCGCCCTGATGCGAGCTTCGCTTGGAGCTGGAAGGTCAAGCATGSCCAGC 960
QY 3059 TGGAGAACAGGCCCTTAAACCTTCCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT 1020
Db 961 TGGAGAACAGGCCCTTAAACCTTCCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT 1020
QY 3119 TCTCTCTGGTGACAGTCTCCCTTGAAACACAGGCTCAGCGTTCAGGGGATTT-- 3176
Db 1021 TCTCTCTGGTGACAGTCTCCCTTGAAACACAGGCTCAGCGTTCAGGGGATTTG 1080
QY 3177 GCCAGCCCCCGGCTCA--CAGTGGGAACACAGGCTCG-- --CAGCAGCAAGGTGGGG 3229
Db 1081 NCCAGCCCCCGGCTCAGAGTGGGAACACAGGCTTCGNCAGCAGCAAGGTGGGG 1140
QY 3230 GCAAGCAATGCTCCAGGATTTTCAACCTGAGCCCTTGCCCAAC-- --CCTGTGA 3283
Db |||||

Db 1141 CAGACNAGATGCTCTCCAGGATTTTCACANCTGAGCCNTGCCCANCCCTGCTGAADA 1200
QY 3284 AAAAACATTCGCGCACGTGAAGAGACAGAGAGATGGC---AGGAGTTACTCTGGGAA 3339
Db 1201 AAACATYTNCCGCGCACGTGAAGAGACAGAGAGATGNCAGGAGTTNNACCTYGGGAA 1260
QY 3340 ACAAAACAGGATCTTT--TTCTGCCCCCTGCTCCAGT-CGAGTTGGCTGA--CCGCTTG 3395
Db 1261 ACAAAACAGGATCTTTNTTCTGCCCCCTGCTCCAGTNCAGTGGCTGNCACCCGCTGG 1320
QY 3396 GATCAGTGACCATTTGTTGGCAGA-CAGGGAGAGCAGCTTCAGCCTGGGTGCAAGGG 3454
Db 1321 ANTCACTGACCATTTGTTGGCAGANACAGGGAGAGCAGCTTCAGCCTGGGTGCAAGGG 1380
QY 3455 GTGGCGAGCCCTTCGCGCCCTCACCCT-CAAGGTGCTGTG-AGAGTGTCAAGTGTGA 3512
Db 1381 GTGGCGAGCCCTTCGCGCCCTCACCCTNCAGGCTGTGTGNAGAGTGTCAAAGTGTGA 1440
QY 3513 AGGGCCCAAACTC--AGGTTTCAGTGACAGAACCAAGGT-CAGCAGGTATGCCGCCGTAG 3568
Db 1441 AGGNCACCAANCTCAGGNTTCAGTGCAAGAACCAAGTNCAGCAGGTATGCCGCCGNTA 1500
QY 3569 GTTAA---GGGGGCCCTTAAACCCCTTGCTTGGCTTCA---CCTGCCAGCTCA-CCC 3620
Db 1501 GGTAAANNNGGGGCCCTCTNAAACCCCTTGCTTNGGCCCTNCACCTINGSCCAGCTCANCC 1560
QY 3621 CTTTTGGGTGTAGGGGAAAGAAATGCTGTGACCTCGGGAAGGT-CCCTGGGTAGAAATAC 3679
Db 1561 CTTTTGGGTGTAGGGGAAAGAAATGCTGTGACCTCGGGAAGGTTCCTTGGTAGAATAC 1620
QY 3680 CACACTTTTCAGTTGTTGCAACAGGTCTGTAGTTGACCTCTGGTTCAGCAGAGGACC 3739
Db 1621 CACACTTTTCAGTTGTTGCAACAGGTCTGTAGTTGACCTCTGGTTCAGCAGAGGACC 1680
QY 3740 AAAGAAGTGTGTAAAGTGAAGTGTCTCAGT-CGCCAGACATGTGCCCTTTGCTGTG 3798
Db 1681 AAAGAAGTGTGTAAAGTGAAGTGTCTCAGTNCACCCAGACATGTGCCCCCTTTGCTGTG 1740
QY 3799 GCTACACTTTCCCGAGCAGCAGCGCCCGAGCCCTTCAGGCCACAGCATGCCCCAG 3858
Db 1741 GCTACACTTTCCCGAGCAGCAGCGCCCGAGCCCTTCAGGCCACAGCATGCCCCAG 1800
QY 3859 ACTCGTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGTAGAGGATATGSCCTGA 3918
Db 1801 ACTCGTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGTAGAGGATATGSCCTGA 1860
QY 3919 CAGGAACAGTCTGTGGATGGAATGATCAGTGTCTAA--GGAAAGCAGCAGAGAGACG- 3975
Db 1861 CAGGAACAGTCTGTGGATGGAATGATCAGTGTCTAAAGGNAAGCAGCAGAGAGAGAGY 1920
QY 3976 TCCGGCGCCCGAGCCCACTATCAGTGT---CAGCGTGTGTTCCCGCAGAGCAGC- 4031
Db 1921 TCCGGCGCCCGAGCCCACTATCAGTGTTCAGTGTTCAGCGTGTGTTCCCGCAGNAGCA 1980
QY 4032 -----TCAGCATCACACTCAGCTACCTCCCTCGCCCTGG--CCAGAGGTTACTG 4082
Db 1981 GCTNCAGNATCANCATGACATNCACCTTNGCCCTCGCCCTGCGCCAGAGGTTACTG 2040
QY 4083 CCG-ACGGCACTTTCAGCTGTG--ATGACCTCAAAGCACTTTCATGCTGCCCTCTGCA 4139
Db 2041 CCGNACGGCACTTTCAGCTGTGACNTCTGATGNACCTCAAAGCACTTTCATGCTGCCCTTNG 2100
QY 4140 GGGCAGGCGCAGGGCAGTGACACT-----GTAGGAGCATAGCAGCCAGGATGG 4189
Db 2101 GCAAGGNCAGGNCAGGNCAGTGACANCTGTAGNAGCATANGCAANGCCAGGAGATGG 2160
QY 4190 GGTGAA--GGGACACAGTCTTGGCTGTCCA-CATGATGTGACT-CCTCAAACTCTT- 4244
Db 2161 GGTGNAAGGNCAGAGTCTTGGCTGTCCANCATGATGATGACTNCCTCAAACTCTTN 2220
QY 4245 --CCAGATTTCTTAAGAAATAGACCCCC--TTCCCAATGCCAGCTTAGCCTTCTCT 4300
Db 2221 NCCAGNATTTCTTAAGAAATAGACCCCTTNGCCCTTNGCCCTTNGCCCTTAGCCTTCTCT 2280

QY 4301 CCCAGGGGAGCTA-CTCAGGACTCACGTAGCATTTAAATCAGCTGTG-AATCGTCAGGGGG 4358
DB 2281 CCCAGGGGAGCTANCTCAGGACTCACGTAGCATTTAAATCAGCTGTGNAATCGTCAGGGGG 2340
QY 4359 TGTCTGCTAGGCTCAACTCTCTGGGCGAGGGAGCGCGAGACTCGGTGGGAGAGCTCAT 4418
DB 2341 TGTCTGCTAGGCTCAACTCTCTGGGCGAGGGAGCGCGAGACTCGGTGGGAGAGCTCAT 2400
QY 4419 TCCACATCTTTGCCAAGACAGCCTTTT-GTCCAGCTGTGCCACATTTGAGTTCAGACTGTCTCCC 4477
DB 2401 TCCACATCTTTGCCAAGACAGCCTTTTGTCTCAGCTGTGCCACATTTGAGTTCAGACTGTCTCCC 2460
QY 4478 GGGGAGAGAGCCCCGGGCCCCCAGACATTAAGAACTGTCAGGCTTTGGTACTGTCAGAGTCTG 4537
DB 2461 GGGGAGAGAGCCCCGGGCCCCCAGACATTAAGAACTGTCAGGCTTTGGTACTGTCAGAGTCTG 2520
QY 4538 GGTGTGTAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGTATGACAAATGTTAAAAAA 4596
DB 2521 GGTGTGTAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGTATGACAAATGTTAAAAAA 2579

RESULT 14

ID ACN44438 standard; DNA; 73967 BP.

XX AC ACN44438;

XX 18-NOV-2004 (first entry)

XX Human genomic sequence hCG27607.

XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

XX Homo sapiens.

XX WO2003073826-A2.

XX 12-SEP-2003.

XX 28-FEB-2003; 2003WO-US006235.

XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-328604/31.

XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma comprises a nucleotide sequence.

XX Claim 1; SEQ ID NO 886; Opp; English.

XX The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published

XX Sequence 73967 BP; 18655 A; 18476 C; 19039 G; 17797 T; 0 U; 0 Other;

XX Query Match 33.9%; Score 1556.2; DB 11; Length 73967;

Best Local Similarity 98.0%; Pred. No. 0; Matches 1650; Conservative 0; Mismatches 23; Indels 10; Gaps 7;
QY 2909 AGATCCAGCTGACAGCCTTTCAGCTTGTGTACCAAGACGGGACGCTGTTTCGTACACACA 2968
DB 62243 AGATCCAGCTGACAGCCTTTCAGCTTGTGTACCAAGACGGGACGCTGTTTCGTACACACA 62302
QY 2969 TGAAGTGTCCAGACTCGGGGATCGACCTGACGTGCAACACTGGCCCCCTGATGGCAGCTTCG 3028
DB 62303 TGAAGTGTCCAGACTCGGGGATCGACCTGACGTGCAACACTGGCCCCCTGATGGCAGCTTCG 62362
QY 3029 CCTGGAGCTCGAGGCTCAAGCATGGCCAGCTGGAGAACAGCCCTAACCCCTGCCCTCCAC 3088
DB 62363 CCTGGAGCTCGAGGCTCAAGCATGGCCAGCTGGAGAACAGCCCTAACCCCTGCCCTCCAC 62422
QY 3089 CGCCGGCTCCACACTGCGGAAAGCAGCCTTCTCTGCTCGGTGACAGATGCTGTCCTGAAAA 3148
DB 62423 CGCCGGCTCCACACTGCGG-AGCAGCCTTCTCTGCTCGGTGACAGATGCTGTCCTG-AA 62480
QY 3149 ACACAGGCTCAGCCGTTCCAGGGGATTTGACAGCCCCCGGCTCACAGTGGGAAACACAGGG 3208
DB 62481 ACACAGGCTCAGCCGTTCCAGGGGATTTGACAGCCCCCGGCTCACAGTGGGAAACACAGGG 62539
QY 3209 CCTCGACAGCAAGGTGGGGGCAAGCAGATGCTCCAGGATTTTACACCTGAGCCCT 3268
DB 62540 CCTCGACAGCAAGGTGGGGGCAAGCAGATGCTCCAGGATTTTACACCTGAGCCCT 62599
QY 3269 GCCCCACCTGCTGAAAAACA-TCCGCCACGTGAAGAGACAGAGGAGATGCGAGGAG 3327
DB 62600 GCCCCACCTGCTGAGAAAAACA-TCCGCCACGTGAAGAGACAG-AGAGGATGCGAGGAG 62658
QY 3328 TTACCTGGGAAAAACAAACAGGGATCTTTTTCTGCCCTGCTCCAGTGCAGTGTGGCCTGA 3387
DB 62659 TTACCTGGGAAAAACAAACAGG-ATCTTCTCTGCTGCTCCAGTGCAGTGTGGCCTGA 62714
QY 3388 CCAGCTTGGATCAGTGACCAATTTGTTGGCAGACAGGGGAGAGAGCTTCCAGCCTGGGTC 3447
DB 62715 CCAGCTTGGATCAGTGACCAATTTGTTGGCAGACAGGGGAGAGAGCTTCCAGCCTGGGTC 62774
QY 3448 AGAAGGGGTGGGAGGCCCTTTCGSCCCTCACCCCTCCAGGCTGCTGTGAGAGTGTCAAGT 3507
DB 62775 AGAAGGGGTGGGAGGCCCTTTCGSCCCTCACCCCTCCAGGCTGCTGTGAGAGTGTCAAGT 62834
QY 3508 GTGTAAGGGGCCAAACTCAGTTCAGTGCAGAAACAGGTCAGCAGGTATGCCCCCGGTA 3567
DB 62835 GTGTAAGGGGCCAAACTCAGTTCAGTGCAGAAACAGGTCAGCAGGTATGCCCCCGGTA 62894
QY 3568 GGTAAAGGGGCCCTCTAAACCCCTTGCCTGGCCCTACCTGGCCAGCTCACCCCTTTGG 3627
DB 62895 GGTAAAGGGGCCCTCTAAACCCCTTGCCTGGCCCTACCTGGCCAGCTCACCCCTTTGG 62954
QY 3628 GTGTAGGGGAAAAAGATGCTGACCTTGGGAAGGCTCCCTGGTAGAATACACACACTTT 3687
DB 62955 GTGTAGGGGAAAAAGATGCTGACCTTGGGAAGGCTCCCTGGTAGAATACACACACTTT 63014
QY 3688 TCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTTCTGTTTCAGCAAGGACCAAGAGG 3747
DB 63015 TCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTTCTGTTTCAGCAAGGACCAAGAGG 63074
QY 3748 TGTGTAAGTGAAGTGTTCAGTTCAGTCCCAGACATGTGCCCTTTGCTGTGGCTTACCACT 3807
DB 63075 TGTGTAAGTGAAGTGTTCAGTTCAGTCCCAGACATGTGCCCTTTGCTGTGGCTTACCACT 63134
QY 3808 CTTCCCCAGAGCAGCAGCCCCCGAGCCCTTTCAGGCCAGCAGCTGCCAGACTGCTGG 3867
DB 63135 CTTCCCCAGAGCAGCAGCCCCCGAGCCCTTTCAGGCCAGCAGCTGCCAGACTGCTGG 63194
QY 3868 CACTCAGTTCCTCATCTGTAAGGTGAAGGGTGTATGACAGGATGCTGACAGGAAACAG 3927
DB 63195 CACTCAGTTCCTCATCTGTAAGGTGAAGGGTGTATGACAGGATGCTGACAGGAAACAG 63254
QY 3928 TCTGTGGATGGACATGATCAGTGTCTAAAGAAACAGCAGAGAGAGAGCG-TCCGGCGCCCC 3986

QY	2548	CTGGAATAGAAATATTCCTCAACAGCCTGTCCCGACCAATTTCTCTGGAGGACGAG	2607
Db	528	CTGGAATAGAAATATTCCTCAACAGCCTGTCCCGACCAATTTCTCTGGAGGACGAG	469
QY	2608	CAAAATTCCTCGTGCCTCAGCATCGACAGCCTCTCCCTGTGGATGACAGTGAGAAGAAC	2667
Db	468	CAAAATTCCTCGTGCCTCAGCATCGACAGCCTCTCCCTGTGGATGACAGTGAGAAGAAC	409
QY	2668	CCATCAAAAGGCCTCTCAAAGCTCGCGGACACCCCTGAGCTCAGGCGTACACTCTCTGGAGC	2727
Db	408	CCATCAAAAGGCCTCTCAAAGCTCGCGGACACCCCTGAGCTCAGGCGTACACTCTCTGGAGC	349
QY	2728	AGCCAGGCGGAGCTCGAAGCTCCAGCTGGAAACATGGTGGCCCGGGGGGGCCCCACC	2787
Db	348	AGCCAGGCGGAGCTCGAAGCTCCAGCTGGAAACATGGTGGCCCGGGGGGGCCCCACC	289
QY	2788	GACACCCCAAGCTATTTCAATGCTGTGAAAGTCCAAATACAGTCTCTTAATGGTGAAACAC	2847
Db	288	GACACCCCAAGCTATTTCAATGCTGTGAAAGTCCAAATACAGTCTCTTAATGGTGAAACAC	229
QY	2848	CTGCACATCCGGGAGTTCCACCGGGTCAAAGTGGGAGACATCGCCACTGGCATCAGCAGC	2907
Db	228	CTGCACATCCGGGAGTTCCACCGGGTCAAAGTGGGAGACATCGCCACTGGCATCAGCAGC	169
QY	2908	CAGATCCAGCTGCAGCCTTCAGCTTGGTCACCAAGCGGCGCCTGTTGCTACGAC	2967
Db	168	CAGATCCAGCTGCAGCCTTCAGCTTGGTCACCAAGCGGCGCCTGTTGCTACGAC	109
QY	2968	ATGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACACTGGCCCCCTGATGGCAGCTTC	3027
Db	108	ATGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACACTGGCCCCCTGATGGCAGCTTC	49
QY	3028	GCCTGAGCTGGAGGTTCAAGCATGGCCAGCTGGAGAACAGGCCCTAA	3075
Db	48	GCCTGAGCTGGAGGTTCAAGCATGGCCAGCTGGAGAACAGGCCCTAA	1

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:20:34 ; Search time 9684.08 Seconds
(without alignments)
18065.048 Million cell updates/sec

Title: US-09-155-676B-6
Perfect score: 4596
Sequence: 1 agcgggggactgtgcctg.....gatgacaaatgttaaaaaa 4596

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gsl1:*

9: gb_gsl2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3309.8	72.0	3971	3	CR749592 Homo sapi
2	2588.4	56.3	2844	9	AY413577 Homo sapi
3	2099.4	45.7	4174	3	AK036200 Mus muscu
4	2095.6	45.6	2765	9	AY413578 Pan trogl
5	1942.2	42.3	2829	9	AY413579 Mus muscu
6	1521.2	33.1	1640	3	CR612342 full-leng
7	943	20.5	1095	1	AL578455 AL578455
8	916	19.9	1133	1	AL556130 AL556130
9	876.2	19.1	1140	4	BM458601 AGENCOURT
10	811.4	17.7	955	5	BU541005 AGENCOURT
11	808.8	17.6	1038	4	BM454415 AGENCOURT
12	787.4	17.1	909	5	BX331085 BX331085
13	756.2	16.5	1037	4	BG775354 AGENCOURT
14	745	16.2	774	4	BG822015 602726122
15	741.6	16.1	2964	3	AK038828 AGENCOURT
16	737.8	16.1	1166	4	BM458535 AGENCOURT
17	714.6	15.5	911	5	BU195099 AGENCOURT
18	705	15.3	762	4	BI259434 602968032
19	684.2	14.9	743	7	CH409541 328775521
20	675	14.7	702	5	EX102204 BX102204
21	665	14.5	741	4	BG685061 602636826
22	658.4	14.3	819	5	BQ231411 AGENCOURT
23	653.8	14.2	896	5	BU542238 AGENCOURT
24	647.8	14.1	707	1	AL042398 DKFZp434B

c 25 635.6 13.8 665 6 CB305346 UI-CF-EN1
26 629.6 13.7 964 5 BQ923823 AGENCOURT
27 622.2 13.5 991 4 BM550076 AGENCOURT
28 605.8 13.2 702 4 BI489907 603031660
29 604.4 13.2 768 4 BM048296 603625505
30 596.2 13.0 806 4 BI161002 602865308
31 591 12.9 701 4 BG774889 602849808
32 580 12.6 702 7 CN409540 170006001
33 568.6 12.4 702 5 CK979910 4111556 B
34 564.2 12.3 580 5 BP267282 BP267282
35 556.2 12.1 579 2 BE407122 601301805
36 549.8 12.0 1172 4 BM543747 AGENCOURT
37 548 11.9 556 2 BE276256 601144348
38 538.4 11.7 795 4 BI826678 603077440
39 530 11.5 736 2 BB632073 BB632073
40 524.4 11.4 548 2 BE391253 601285947
41 520 11.3 531 6 CA748591 UI-H-F11-
42 518.8 11.3 659 2 AM178916 MRO-ST001
43 512.6 11.2 556 2 BE388034 601283733
44 512.2 11.1 702 4 BI554785 603236762
45 512.2 11.1 730 1 AI956650 ul79h03.y

ALIGNMENTS

RESULT 1
LOCUS CR749592 3971 bp mRNA linear HTC 19-AUG-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686J04131 (from clone DKFZp686J04131).
ACCESSION CR749592
VERSION CR749592.1 GI:51476844
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3971)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Bloembergen, H., Boecher, M., Brandt, P., Meves, H.W., Weill, B., Amid, C.,
Osanger, A., Fobo, G., Han, M. and Wiemann, S.
CONSRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686J04131) is available at the RZPD Deutsches
Reisourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686J04131
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES

source
1..3971
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686J04131"
/tissue_type="uterus"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/note="MAP3K14 protein, N-terminus truncated,
differentially spliced"
1..3971
/gene="DKFZp686J04131"
/gene="DKFZp686J04131"
/codon_start=1
/product="hypothetical protein"

/protein_id="CAH18391.1"
/db_xref="GI:51476845"
/translation="CFKOLGELRPALPRSELHKLISPLQCLINHWKLLHPDQGGPLP
LPHPPPHPPLQPKPPLPESPLGKLACVDQKPLPDLPHLSKLACVDSPKPLPPLP
PSCUSGAHEKFPSEYILVHALQGSVSSQAHSLTSIAKTWAARSGRSPKPTEDN
EGLVLETKLKPVDYREBEVHAWTHQLRLGRSGFEVHRMEDKQTFQCAVKKRVLEV
FRAEELMACAGLTSPRIYVPLGAVREGPWNI FMBLEGGSLGLVKEQGLPEDRAL
YVLGALRGLEYLHSHRRILIHGDVKADNVLLSDGSHAALCDFGHAVCLIQDGLGKSLI
TGDXTIPGTEHMAPRVILGRSCDAKVDVWSSCCMWLHMLNGCHPWQTFRRGPLCLKVG
DGTQHGGQBELQGPQAQAGRWGCSAVADGARGGLVIVHGS"

ORIGIN

Query Match 72.0%; Score 3309.8; DB 3; Length 3971;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 3673; Conservative 0; Mismatches 32; Indels 244; Gaps 10;
QY 856 TGTTTTAGCAGCTTGGGAGGGGCTACGGCCGGCTCTGCCCTCGATCAGAACTCCACAAA 915
Db 1 TGTTTTAAGCAGCTTGGGAGGGGCTACGGCCGGCTCTGCCCTCGATCAGAACTCCACAAA 60
QY 916 CTGATCAGCCCTTCAATGCTGAACACAGTGTGAAACTGCACACCCCGCAGGACGGA 975
Db 61 CTGATCAGCCCTTCAATGCTGAACACAGTGTGAAACTGCACACCCCGCAGGACGGA 120
QY 976 GGCCCCCTGCCCTGCCACGACGCCCTTCCCTATAGCAGACTGCCTCATCCCTTCCCA 1035
Db 121 GGCCCCCTGCCCTGCCACGAC-----CCCTTCCCA 153
QY 1036 TTCCACCTCTCAGCCCTGGAACCTCACTCTGAGTCTCTTCTGGGCAAACTGGCC 1095
Db 154 TTCCACCTCTCAGCCCTGGAACCTCACTCTGAGTCTCTTCTGGGCAAACTGGCC 213
QY 1096 TGTGTAGACAGCAGAAACCTTGCCTGACCCACACCTGAGCAAACTGSCCTGTGTAGAC 1155
Db 214 TGTGTAGACAGCAGAAACCTTGCCTGACCCACACCTGAGCAAACTGSCCTGTGTAGAC 273
QY 1156 AGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGTCTGTGGTGGCCAT 1215
Db 274 AGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGTCTGTGGTGGCCAT 333
QY 1216 GAGAGTTTCTGTGGAGAAATACCTAGTGCATGTCTTGCAAGGAGCGTGTAGCTCAAGC 1275
Db 334 GAGAGTTTCTGTGGAGAAATACCTAGTGCATGTCTTGCAAGGAGCGTGTAGCTCAGGC 393
QY 1276 CAGGCCCAAGCTCAGCAGCTGCGCAAGACCTGGGAGCAGCGGGCTCCAGATCCCGG 1335
Db 394 CAGGCCCAAGCTCAGCAGCTGCGCAAGACCTGGGAGCAGCGGGCTCCAGATCCCGG 453
QY 1336 GAGCCAGCCCCAAACTGAGGACAAACGAGGGTGTCTGCTCACTGAGAAACTCAAGCCA 1395
Db 454 GAGCCAGCCCCAAACTGAGGACAAACGAGGGTGTCTGCTCACTGAGAAACTCAAGCCA 513
QY 1396 GTGGATTATGATACGAGAAAGTTCATCTGGGCCACGCACAGCTCGCCTGGGGCAGA 1455
Db 514 GTGGATTATGATACGAGAAAGTTCATCTGGGCCACGCACAGCTCGCCTGGGGCAGA 573
QY 1456 GGCTCTTTCGAGAGTGCACAGGATGAGGACAGCAGCTGCTTCCAGTCCGCTGTC 1515
Db 574 GGCTCTTTCGAGAGTGCACAGGATGAGGACAGCAGCTGCTTCCAGTCCGCTGTC 633
QY 1516 AAAAAGGTGCGCTGGAAGTATTTTCGGGACAGAGGAGCTGATGGCATGTGCAGGATTGACC 1575
Db 634 AAAAAGGTGCGCTGGAAGTATTTTCGGGACAGAGGAGCTGATGGCATGTGCAGGATTGACC 693
QY 1576 TCACCAGAAATGTCCCTTTGTATGGAGCTGTGAGAGAGGGCTTTGGGTCAAACTTTTC 1635
Db 694 TCACCAGAAATGTCCCTTTGTATGGAGCTGTGAGAGAGGGCTTTGGGTCAAACTTTTC 753
QY 1636 ATGGAGCTCTCGAAGGTGGCTCCCTGGCCAGAGCTGTTCAAGGACAGGGCTGTCTCCCA 1695
Db 754 ATGGAGCTCTCGAAGGTGGCTCCCTGGCCAGAGCTGTTCAAGGACAGGGCTGTCTCCCA 813
QY 1696 GAGGACCGGGCCCTGTGTAACCTGGGCCAGGCCCTGGAGGGTCTGGAAATACCTCACTCA 1755

Db 814 GAGGACCGGGCCCTGTACTACTCTGGCCAGGCCCTGGAGGGTCTGGAAATACCTCCACTCA 873
QY 1756 CGAAGGATTTCTGCATGGGAGCGTCAAAAGCTCACAACGTCCTCTGTCTCAGCGATGGGAGC 1815
Db 874 CGAAGGATTTCTGCATGGGAGCGTCAAAAGCTCACAACGTCCTCTGTCTCAGCGATGGGAGC 933
QY 1816 CACGAGCCCTCTGTGACTTTTGGGCGCATGCTGTGTCTTCAACCTGTATGGCTTGGGAAAG 1875
Db 934 CACGAGCCCTCTGTGACTTTTGGGCGATGCTGTGTCTTCAACCTGTATGGCTTGGGAAAG 993
QY 1876 TCCTTGTCTCAGGGGACTACATCTCTGGCACAGAGACCCACATAGCTTCCGAGAGTGTGTG 1935
Db 994 TCCTTGTCTCAGGGGACTACATCTCTGGCACAGAGACCCACATGGCTCCGAGAGTGTGTG 1053
QY 1936 CTGGCAGGAGCTCGGACGCCAAGTGGATGCTCTGGAGCAGTGTGATGATGCTGCAC 1995
Db 1054 CTGGCAGGAGCTCGGACGCCAAGTGGATGCTCTGGAGCAGTGTGATGATGCTGCAC 1113
QY 1996 ATGTCTAACGGCTGCCACCCCTTGGACTCAGTTCCTTCCGAGGGCGCTCTGCCTCA 2050
Db 1114 ATGTCTAACGGCTGCCACCCCTTGGACTCAGTTCCTTCCGAGGGCGCTCTGCCTCAAGTTC 1173
QY 2051 ----- 2050
Db 1174 GGTGACGGCACCCAGCATGGCCAAACAGGGGGAGCTGCAGGGCCCTGGGGCCAGGCTGCA 1233
QY 2051 ----- 2050
Db 1234 GGCTGGAGGGCTGCTCTGCGGTGGCAGATGGAGCAAGGGAGGCTGTGTATTCATGCTGTG 1293
QY 2051 ----- 2050
Db 1294 GGATCTTAATTCAAATACTTCCACCCCAATCTGGGGGTACAAGGGGAGATCCAGCTGG 1353
QY 2051 -----AGATTGCCAGGAGCTCTCCGCTGTGTAGGGAGATCCAC 2089
Db 1354 GAGGCTCACCAAGTTTCATGGCAGATTGTCAGCAGGCTCTCCGCTGTGTAGGGAGATCCAC 1413
QY 2090 CCTCTGCGCCCTCTCACAGCCAGGCCATCCAGAGGGGTGAGGAAAGAGCCCATCC 2149
Db 1414 CCTCTGCGCCCTCTCACAGCCAGGCCATCCAGAGGGGTGAGGAAAGAGCCCATCC 1473
QY 2150 ACCGCGTCTCTGACGCGAGCTGGGAGGGAAGTGAACCGGGGCACTACAGCAAGTGGAG 2209
Db 1474 ACCGCGTCTCTGACGCGAGCTGGGAGGGAAGTGAACCGGGGCACTACAGCAAGTGGAG 1533
QY 2210 GTCTGAAGAGCCCTTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACGCCAAATCAAG 2269
Db 1534 GTCTGAAGAGCCCTTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACGCCAAATCAAG 1593
QY 2270 CCNATTAACCAAGACCCCTCCATGCCAGCCGAGAGAGCTTTTCGCAAGGGGCCCCAGGGC 2329
Db 1594 CCNATTAACCAAGACCCCTCCATGCCAGCCGAGAGAGCTTTTCGCAAGGGGCCCCAGGGC 1653
QY 2330 CCGGCCAGCTGAGGAGACAAACAGCAGAGCCCTTAAGCTCCAGGCTCTCTCTCCACCCAG 2389
Db 1654 CCGGCCAGCTGAGGAGACAAACAGCAGAGCCCTTAAGCTCCAGGCTCTCTCTCCACCCAG 1713
QY 2390 AGCCCCCAGAGCCAAACAAAGTCTCTCTCTTGAATTTTTCAGCAAGGAGGAGTCTGGGATGT 2449
Db 1714 AGCCCCCAGAGCCAAACAAAGTCTCTCTCTTGAATTTTTCAGCAAGGAGGAGTCTGGGATGT 1773
QY 2450 GGGAAACCTTACCTCTGTCTCTCTCTGGAGCAGCCCTCCGCAAGAACCCAGCTCACCAG 2509
Db 1774 GGGAAACCTTACCTCTGTCTCTCTCTGGAGCAGAACCCCTTCGCAAGAACCCAGCTCACCAG 1833
QY 2510 AGCGAAAGAACCCCTCCCGGAGCAGGAACTGTCAGCAGCTGGAATAGAAATATTATTCCTCA 2569
Db 1834 AGCGAAAGAACCCCTCCCGGAGCAGGAACTGTCAGCAGCTGGAATAGAAATATTATTCCTCA 1893
QY 2570 ACAGCCTGTCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTCGCTCAGCA 2629

Db 1894 ACAGCTGTCCAGCCATTTTCTCTGAGGAGCAGGAACAAATTCTCTCGTGCCTCAACA 1953
Qy 2630 TCAGACAGCTCTCCCTGTCGATGACAGTGAGAGAACCCATCAAGGCCCTCTCAAGACT 2689
Db 1954 TCAGACAGCTCTCCCTGTCGATGACAGTGAGAGAACCCATCAAGGCCCTCTCAAGACT 2013
Qy 2690 CGGGGACACCTGAGCTAGCTAGCGGTACACTCTGAGGAGCAGGCGGAGCTCGAAGCT 2749
Db 2014 CGGGGACACCTGAGCTAGCGGTACACTCTGAGGAGCAGGCGGAGCTCGAAGCC 2073
Qy 2750 CCAGCTGAACATAGTGTGCTGCGCGGGCGGCCACCGACACCCCAAGCTATTTCATG 2809
Db 2074 CCAGCTGAACATAGTGTGCTGCGCGGGCGGCCACCGACACCCCAAGCTATTTCATG 2133
Qy 2810 GTGTGAAGTCCAAATACAGTCTCTTAATGTGTGAACACTCTGACATCCGGGAGTTCAC 2869
Db 2134 GTGTGAAGTCCAAATACAGTCTCTTAATGTGTGAACACTCTGACATCCGGGAGTTCAC 2193
Qy 2870 GGGTCAAGTGGGAGACATCGCCACTTGGCATCAGGAGCAGATCCGAGCTGCAGCCTTCA 2929
Db 2194 GGGTCAAGTGGGAGACATCGCCACTTGGCATCAGGAGCAGATCCGAGCTGCAGCCTTCA 2253
Qy 2930 GCTTGGTACCAAGACGGGAGCCTGTTCGTACGATGAGTGCAGACTCGGCA 2989
Db 2254 GCTTGGTACCAAGACGGGAGCCTGTTCGTACGATGAGTGCAGACTCGGCA 2313
Qy 2990 TCAGCTGCACTGACACACTGCGCCCTGATGCGAGCTTCGCTGAGGCTGAGGCTCAAGC 3049
Db 2314 TCAGCTGCACTGACACTGCG-CCCTGATGCGAGCTTCGCTGAGGCTGAGGCTCAAGC 2372
Qy 3050 ATGCGCAGCTGGAGAACAGGCGCTTAACCTCGCTTCCACCGCGCTCCACACTGCGCGA 3109
Db 2373 ATGCGCAGCTGGAGAACAGGCGCTTAACCTCGCTTCCACCGCGCTCCACACTGCGCG- 2431
Qy 3110 AAGCAGCTTCTGCTCGGTGACAGTCTGCTGAAACACAGGCTCAGCCTTCCCA 3169
Db 2432 AAGCAGCTTCTGCTCGGTGACAGTCTGCTG-AAAACAGGCTCAGCCTTCCCA 2490
Qy 3170 GGGGATTGCCAGCCCCCGGCTCAGTGGGAACAGGCGCTCGCAGCAGCAAGTGGG 3229
Db 2491 GGGATCTGCCAG-CCCCGGCTCAGCAGTGGGACAGGCGCTCGCAGCAGCAAGTGGG 2549
Qy 3230 GCAGCAGATGCTCCAGGATTTCACCTGAGCCTGCCCACTGCTGTAAGAAC 3289
Db 2550 GCAGCAGATGCTCCAGGATTTCACCTGAGCCTGCCCACTGCTGTAAGAAC 2609
Qy 3290 A-TCCGCCACTGAAGACAGAGAGGAGTTCAGTGGGATTCCTGGGAAACAAACAG 3348
Db 2610 ACTCCGCCACTGAAGACAG-AGGAGGATGGCAGGATTCCTCGGAAACAAACAG 2668
Qy 3349 GGATCTTTTCTGCCCCCTCCAGTCCAGTGGCTGACCGCTTGGATCAGTGACCAT 3408
Db 2669 ----ATCTTCTGCCCCCTCCAGTCCAGTGGCTGACCGCTTGGATCAGTGACCAT 2724
Qy 3409 TTGTTGGCAGACAGGGAGCAGCTTCAGCCTGGGTGAGAGGGGTGGCGAGCCCTT 3468
Db 2725 TTGTTGGCAGACAGGGAGCAGCTTCAGCCTGGGTGAGAGGGGTGGCGAGCCCTT 2784
Qy 3469 CGGCCCTCACCCTCCAGGCTGCTGAGAGTGTCAAGTGTGAAGGCCCAACTCAGG 3528
Db 2785 CGGCCCTCACCCTCCAGGCTGCTGAGAGTGTCAAGTGTGAAGGCCCAACTCAGG 2844
Qy 3529 TTCAGTCAGAACAGGTCAGCAGTATGCGCGCCGCTGAGGTAAAGGGGCGCTCAAC 3588
Db 2845 TTCAGTCAGAACAGGTCAGCAGTATGCGCGCCGCTGAGGTAAAGGGGCGCTCAAC 2904
Qy 3589 CCCTTGCCTGCCCTCACCCTGCCAGCTCACCCCTTTTGGGTGAGGGGAAAGATGCT 3648
Db 2905 CCCTTGCCTGCCCTCACCCTGCCAGCTCACCCCTTTTGGGTGAGGGGAAAGATGCT 2964
Qy 3649 GACCTGGGAAGGCTCCCTGTTAGAAATACACACTTTTCAGTGTGTGCAACACAGGT 3708
Db 2965 GACCTGGGAAGGCTCCCTGTTAGAAATACACACTTTTCAGTGTGTGCAACACAGGT 3024

Qy 3709 CTTGAGTTGACCTCTCTGTTTCCAGCCAGGACAAAGAGGTGTGTAAAGTGAAGTGTCTC 3768
Db 3025 CTTGAGTTGACCTCTCTGTTTCCAGCCAGGACAAAGAGGTGTGTAAAGTGAAGTGTCTC 3084
Qy 3769 AGTCCCCCAGACATGTGCCCCCTTTGCTGCTGGCTACACTCTTCCCCAGAGCAGCAGGCC 3828
Db 3085 AGTCCCCCAGACATGTGCCCCCTTTGCTGCTGGCTACCACTCTTCCCCAGAGCAGCAGGCC 3144
Qy 3829 CGAGCCCCCTTCCAGCCCCCAGACCTGCCCCCAGACTCGCTGGCAGCTCAGTTCCTCATCTGTA 3888
Db 3145 CGAGCCCCCTTCCAGCCCCCAGACCTGCCCCCAGACTCGCTGGCAGCTCAGTTCCTCATCTGTA 3204
Qy 3889 AAGTGAAGGCTGATGAGGATATGCTGACAGAAACAGTCTGTGGATGAGCATGATCAG 3948
Db 3205 AAGTGAAGGCTGATGAGGATATGCTGACAGAAACAGTCTGTGGATGAGCATGATCAG 3264
Qy 3949 TGCTAAGGAAGCAGCAGAGAGAGAG-TCGGGGCCCCCAGCCCACTATCAGTGTCCAG 4007
Db 3265 TGCTAAGGAAGCAGCAGAGAGAGAGAGCTCGGGCCCCCAGCCCACTATCAGTGTCCAG 3324
Qy 4008 CGTCTGTTTCCCCCAGAGCAGCTCAGCATCACACTGACACTCACCTGCCCCCTGCCCT 4067
Db 3325 CGTCTGTTTCCCCCAGAGCAGCTCAGCATCACACTGACACTCACCTGCCCCCTGCCCT 3384
Qy 4068 GGGCAGAGGCTACTGCGCAGCGCACTTTGCACTCTGATGACCTCAAGCACTTTTCAATGCC 4127
Db 3385 GGGCAGAGGCTACTGCGCAGCGCACTTTGCACTCTGATGACCTCAAGCACTTTTCAATGCC 3444
Qy 4128 TGCCCTCTGCGAGGCGAGGCGAGTGCACACTGAGGAGTGCAGGAGTGCAGGAGAT 4187
Db 3445 TGCCCTCTGCGAGGCGAGGCGAGTGCACACTGAGGAGTGCAGGAGTGCAGGAGAT 3504
Qy 4188 GGGTGAAGGACACAGTCTTGGCTGCTCCACATGCTGACCTCAAACTCTTCCA 4247
Db 3505 GGGTGAAGGACACAGTCTTGGCTGCTCCACATGCTGACCTCAAACTCTTCCA 3564
Qy 4248 GATTCTCTAAGAAATAGCACCCCTTCCCATTTGCCCAGCTTAGCCTTCTTCCCAAGG 4307
Db 3565 GATTCTCTAAGAAATAGCACCCCTTCCCATTTGCCCAGCTTAGCCTTCTTCCCAAGG 3624
Qy 4308 GAGTCTCAGGACTCAGTGAATTAATCAGCTGTGAATCTGTCAGGGGTGTCTGCTA 4367
Db 3625 GAGTCTCAGGACTCAGTGAATTAATCAGCTGTGAATCTGTCAGGGGTGTCTGCTA 3684
Qy 4368 GCCTCAACCTCTGGGCGAGGAGCGCGAGACTCCGTTGGGAGAGTCTATTCACACATC 4427
Db 3685 GCCTCAACCTCTGGGCGAGGAGCGCGAGACTCCGTTGGGAGAGTCTATTCACACATC 3744
Qy 4428 TTGCCAAGACAGCCTTTTGTCCAGCTGTCCACATTTGAGTTCAGACTGCTCCCGGGGAGAG 4487
Db 3745 TTGCCAAGACAGCCTTTTGTCCAGCTGTCCACATTTGAGTTCAGACTGCTCCCGGGGAGAG 3804
Qy 4488 CCGCGGCCCGCAGCACAATAAGAACTCAGCCTTGGTACTGTCAGAGTCTGGGTGTAGAG 4547
Db 3805 CCGCGGCCCGCAGCACAATAAGAACTCAGCCTTGGTACTGTCAGAGTCTGGGTGTAGAG 3864
Qy 4548 AACTCTTTTGAAGCAATAAGTTTGGGTGATGACAAATGTAAAAA 4596
Db 3865 AACTCTTTTGAAGCAATAAGTTTGGGTGATGACAAATGTAAAAA 3913

RESULT 2

AY413577
LOCUS
DEFINITION Homo sapiens MAP3K14 gene, VIRTUAL TRANSCRIPT, partial sequence.
ACCESSION AY413577
VERSION AY413577.1 GI:39769539
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 2844)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
2 (bases 1 to 2844)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source
1..2844
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene
1..2844
/gene="WAP3K14"
/locus_tag="HCM4937"
ORIGIN
Query Match 56.3%; Score 2588.4; DB 9; Length 2844;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 2592; Conservative 0; Mismatches 252; Indels 0; Gaps 0;
QY 232 ATGGCAGTGATGGAATGCGCTGCCAGGTGCCCCCTGCTCAGCAGTGGGGCAGCAGAAG 291
DB 1 ATGGCAGTGATGGAATGCGCTGCCAGGTGCCCCCTGCTCAGCAGTGGGGCAGCAGAAG 60
QY 292 GAACATCCCAAGCCAAAGAGAGAGCGCCGACCTGGGGAAGAAACAGAGCTCCCTCTAC 351
DB 61 GAACATCCCAAGCCAAAGAGAGAGCGCCGACCTGGGGAAGAAACAGAGCTCCCTCTAC 120
QY 352 AAGCTTGAGCGCGTGAGAGAGCCCTGTGTTCTCGGAAAGTGGAGATCTCTGAATGAC 411
DB 121 AAGCTTGAGCGCGTGAGAGAGCCCTGTGTTCTCGGAAAGTGGAGATCTCTGAATGAC 180
QY 412 GTGATTACCAAGGGCAGACCCCAAGAGAGGCTCCGAGGCGAGGCGCAGCTGCATCTATC 471
DB 181 GTGATTACCAAGGGCAGACCCCAAGAGAGGCTCCGAGGCGAGGCGCAGCTGCATCTATC 240
QY 472 ATCGCCCGAGCTGAGTGTGAGATAGCCAGAGTTCAGCCCAACCTTTTCAGAACGCATT 531
DB 241 ATCGCCCGAGCTGAGTGTGAGATAGCCAGAGTTCAGCCCAACCTTTTCAGAACGCATT 300
QY 532 TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACCAAT 591
DB 301 TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACCAAT 360
QY 592 GTGGGCCATGTCTACAGAGGCGAAATGGCCCGTGTGTGTTGGAAGGGAAGCGTCGCAGC 651
DB 361 GTGGGCCATGTCTACAGAGGCGAAATGGCCCGTGTGTGTTGGAAGGGAAGCGTCGCAGC 420
QY 652 AAAGCCCGAAGAACGGAAGAGAGAGCTCAAGTCCCTGGCTCATGCGAGGTGGCC 711
DB 421 AAAGCCCGAAGAACGGAAGAGAGAGCTCAAGTCCCTGGCTCATGCGAGGTGGCC 480
QY 712 TTGGCCAAACCCCTCCCAAGGACCCCTGAGCAGGAGAGCTGCACCATCCCAAGTGCAGGAG 771
DB 481 TTGGCCAAACCCCTCCCAAGGACCCCTGAGCAGGAGAGCTGCACCATCCCAAGTGCAGGAG 540
QY 772 GATAGTCTCCACTCGGCGCCCCATATGTTAGAAACACCCCGAGTTCACCAAGCCTCTG 831
DB 541 GATAGTCTCCACTCGGCGCCCCATATGTTAGAAACACCCCGAGTTCACCAAGCCTCTG 600

QY 832 AAGGAACAGGCGCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGCTACGGCGGCT 891
DB 601 AAGGAACAGGCGCTTGGGNN 660
QY 892 CTGCGCTGCATCAGAACTCCAAACTGATCAGCCCTTGGCAATGTCTGAAACACACCTGTGG 951
DB 661 NNN 720
QY 952 AAATGTCACACCCCGGAGCGGAGGCGCCCTGCCCTGCCACGCAACCCCTTCCCTAT 1011
DB 721 AAATGTCACACCCCGGAGCGGAGGCGCCCTGCCCTGCCACGCAACCCCTTCCCTAT 780
QY 1012 AGCAGACTGCTCATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAACCTCACCTCTG 1071
DB 781 AGCAGACTGCTCATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAACCTCACCTCTG 840
QY 1072 GAGTCTTCTGGGCAAACTGGCTGTGTAGACAGCCAGAAACCTTCCCTGACCCACAC 1131
DB 841 GAGTCTTCTGGGCAAACTGGCTGTGTAGACAGCCAGAAACCTTGCCTGACCCACAC 900
QY 1132 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTGCCCTGGGCCACACCTGGAGCCC 1191
DB 901 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTGCCCTGGGCCACACCTGGAGCCC 960
QY 1192 AGCTGCTGTCTCGTGGTCCCATGAGAAATTTCTGTGGAGGAATACCTAGTGCATGCT 1251
DB 961 AGCTGCTGTCTCGTGGTCCCATGAGAAATTTCTGTGGAGGAATACCTAGTGCATGCT 1020
QY 1252 CTGCAAGCAGCGTGAGCTCAAGCCAGCCACAGCTGACAGCTGGCCAAAGACCTGG 1311
DB 1021 CTGCAAGCAGCGTGAGCTCAAGCCAGCCACAGCTGACAGCTGGCCAAAGACCTGG 1080
QY 1312 GCAGCACGGGCTCCAGATCCCGGAGCCAGCCCAAACTGAGAGCAACGAGGGTGC 1371
DB 1081 GCAGCAAGGGGCTCCAGATCCCGGAGCCAGCCCAAACTGAGAGCAACGAGGGTGC 1140
QY 1372 CTGCTCACTGAGAACTCAAGCCAGTGATTTAGTACCGAGAAAGTCCACCTGGGCC 1431
DB 1141 CTGCTCACTGAGAACTCAAGCCAGTGATTTAGTACCGAGAAAGTCCACCTGGGCC 1200
QY 1432 AGCACCAGCTCCGCTGGGAGAGGCTCTTGGAGAGGTGCACAGATGGAGACAAG 1491
DB 1201 AGCACCAGCTCCGCTGGGAGAGGCTCTTGGAGAGGTGCACAGATGGAGACAAG 1260
QY 1492 CAGATGCTTCCAGTGCCTGTCAAAAAGTGCCTGGAAGTATTTGGGCGAGAGAG 1551
DB 1261 CAGATGCTTCCAGTGCCTGTCAAAAAGTGCCTGGAAGTATTTGGGCGAGAGAG 1320
QY 1552 CTGATGGCATGTGCAAGGATTCACCTCAGCAATTTGTCCTTTGTATGGAGCTGTGAGA 1611
DB 1321 CTGATGGCATGTGCAAGGATTCACCTCAGCAATTTGTCCTTTGTATGGAGCTGTGAGA 1380
QY 1612 GAAGGCGCTTGGGTCAACATCTTCATGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1671
DB 1381 GAAGGCGCTTGGGTCAACATCTTCATGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1440
QY 1672 GTCAAGGAGCAGGCTGTCTCCAGAGACCGGGCCCTGTACTACTGGGCCAGGCCCTG 1731
DB 1441 GTCAAGGAGCAGGCTGTCTCCAGAGACCGGGCCCTGTACTACTGGGCCAGGCCCTG 1500
QY 1732 GAGGCTCTGGAATACCTCAGTCAAGAGGATTTGTCATGGGAGCGTCAAGCTCAGCAAC 1791
DB 1501 GAGGCTCTGGAATACCTCAGTCAAGAGGATTTGTCATGGGAGCGTCAAGCTCAGCAAC 1560
QY 1792 GTGCTCTGTCCAGCATGGGAGCAGCAGAGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1851
DB 1561 GTGCTCTGTCCAGCATGGGAGCAGCAGAGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1620
QY 1852 CTTCAACCTGATGGCTGGAAAGTCTTGTGTACAGGGGACTACATCCCTGGGCACAGAG 1911
DB 1621 CTTCAACCTGATGGCTGGAAAGTCTTGTGTACAGGGGACTACATCCCTGGGCACAGAG 1680

QY 1912 ACCACATGGCTCCGAGGTGGTGTGGGAGGAGCTGCGAGCGCAAGGTGGATGTTCTGG 1971
Db 1681 ACCACATGGCTCCGAGGTGGTGTGGGAGGAGCTGCGAGCGCAAGGTGGATGTTCTGG 1740
QY 1972 AGCAGCTCTGTATGATGTCACATGCTCAACGCTGCGACCCCTGAGCTTCATGTTCTTC 2031
Db 1741 AGCAGCTCTGTATGATGTCACATGCTCAACGCTGCGACCCCTGAGCTTCATGTTCTTC 1800
QY 2032 CGAGGGCGCTCTGCTCAAGATTGCCAGCGAGCCTCGCCCTGTGAGGGAGATCCCAACC 2091
Db 1801 CGAGGGCGCTCTGCTCAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1860
QY 2092 TCCTGCGCCCTCTCACAGCCAGCCATCAAGAGGGCTGAGAAAGAGCCCATCCAC 2151
Db 1861 NNN 1920
QY 2152 CGCGTGTCTGAGCGGAGCTGGAGGGAAGGTGAACCGGCACTACAGCAAGTGGAGGT 2211
Db 1921 NNN 1980
QY 2212 CTGAAGAGCCCTTGGAGGGGAGAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCC 2271
Db 1981 CTGAAGAGCCCTTGGAGGGGAGAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCC 2040
QY 2272 AATTACCAACAGACCTCCATGCCAGCCGAGAGAGCTTTGCCAAGGGCCCAAGGGCC 2331
Db 2041 AATTACCAACAGACCTCCATGCCAGCCGAGAGAGCTTTGCCAAGGGCCCAAGGGCC 2100
QY 2332 CGGCCAGCTGAGGAGACACAGGAGGAGCCCTAAGCTCAGCTCCTCTCCACCCAGAG 2391
Db 2101 CGGCCAGCTGAGGAGACACAGGAGGAGCCCTAAGCTCAGCTCCTCTCCACCCAGAG 2160
QY 2392 CCCCAGAGCCAAACAGTCTCTCCCTTGACTTTGACCAAGGAGGAGTCTGGGATGTGG 2451
Db 2161 CCCCAGAGCCAAACAGTCTCTCCCTTGACTTTGACCAAGGAGGAGTCTGGGATGTGG 2220
QY 2452 GAACCTTACCTCTGTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACAGAG 2511
Db 2221 GAACCTTACCTCTGTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACAGAG 2280
QY 2512 CGGAAGCAACCGTCCGAGCAGGACTGAGCAGCTGGAATAGATTTATTCCTCAAC 2571
Db 2281 CGGAAGCAACCGTCCGAGCAGGAACTGAGCAGCTGGAATAGATTTATTCCTCAAC 2340
QY 2572 AGCTGTCTCCAGCCATTTCTCTGAGGAGCAGGACAAATTTCTCTGCTCTCAGCATC 2631
Db 2341 AGCTGTCTCCAGCCATTTCTCTGAGGAGCAGGACAAATTTCTCTGCTCTCAGCATC 2400
QY 2632 GACAGCTCTCTCTCGATGACAGTGAGAGAACCCATCAAGGGCTCTCAAGCTCG 2691
Db 2401 GACAGCTCTCTCTCGATGACAGTGAGAGAACCCATCAAGGGCTCTCAAGCTCG 2460
QY 2692 CGGACACCTCTGAGCTCAGGGGTACATCTCTGAGAGCAGCGCCGAGGCTCGAAGCTCC 2751
Db 2461 CGGACACCTCTGAGCTCAGGGGTACATCTCTGAGAGCAGCGCGGAGGCTCGAAGCTCC 2520
QY 2752 AGCTGGAACATGGTGTGGCCCGGGGGCGCCACCGACACCCCAAGCTATTTCAATGGT 2811
Db 2521 AGCTGGAACATGGTGTGGCCCGGGGGCGCCACCGACACCCCAAGCTATTTCAATGGT 2580
QY 2812 GTGAAGTCCAAATACAGTCTCTTAATGGTGAACACCTGCAATCCGGAGTTCACCGG 2871
Db 2581 GTGAAGTCCAAATACAGTCTCTTAATGGTGAACACCTGCAATCCGGAGTTCACCGG 2640
QY 2872 GTCAAGTGGAGACATCGCCACTGCGCATCAGCAGCCAGATCCAGCTGCAAGCTTCAGC 2931
Db 2641 GTCAAGTGGAGACATCGCCACTGCGCATCAGCAGCCAGATCCAGCTGCAAGCTTCAGC 2700
QY 2932 TTGGTCCAAAGAGCGGCGCTGTTCCGTACGACATGGAGGTCAGAGCTCGGGCATC 2991
Db 2701 TTGGTCCAAAGAGCGGCGCTGTTCCGTACGACATGGAGGTCAGAGCTCGGGCATC 2760
QY 2992 GACCTGAGTGCACATGGGCCCTCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAGAGCAT 3051

Db 2761 GACCTGAGTGCACATGGGCCCTCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAGAGCAT 2820
QY 3052 GGCCAGCTGGAGACACAGCCCTTAA 3075
Db 2821 GGCCAGCTGGAGACACAGCCCTTAA 2844
RESULT 3
AK036200
LOCUS
DEFINITION
Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:9630045G21 product:mitogen-activated
protein kinase kinase 14, full insert sequence.
AK036200
VERSION
AK036200.1 GI:26331227
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayaashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayaashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED
11042159
REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayaashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED
11076861
REFERENCE
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
REFERENCE
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
REFERENCE
6
(bases 1 to 4174)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayaashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sugabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayaashizaki, Y.
Direct Submission
TITLE

JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://fantom.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

FEATURES

Location/Qualifiers

source

1..4174

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:9630045G21"

/db_xref="taxon:10090"

/clone="9630045G21"

/tissue_type="cerebellum"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="16 days neonate"

85..2913

/note="unnamed protein product; mitogen-activated protein

kinase kinase kinase 14 (MGD|MGI:1858204, GB|NM_016896,

evidence: BLASTN, 100%, match=2829)

putative"

/codon_start=1

/protein_id="BAC29144.1"

/db_xref="GI:26331228"

/translation="MAWMEVACPPTPGSANGQOKELAKAKTKTSLGKQKSCIFKLEA

VEKSPVFCQKILNDVITKADGSEGGPPAISIIAACEENSQEFSPFSEIRFI

AGSQYVSSESIDQIPNNVAHATEGMARVCRGRHGRKRRKRSKSLAQAGVA

LAKPLRPTPEQSSCTIPVOEDSEPLNLYARNVSQFTKPLGGPGLHLGKQDEGLR

LVLPRLPELHKLISPLQCNLHVWLHPQATGPRPHPTFPFVSGMHPFPFVPLEPWK

PYMLDSALVDLKGAVGQRLPPLPPLHLSQLAHGDQKPLPGLHSSCPSSGALEKVP

VEEYLVAHQSGVSGQSHSLASLAKTWSSGSQAKLQRLGPTEDENEGVELLEKLPVD

YBIEVEVHMTHPQRVSGFGEVHRMDKQTFQCAVKKRVLEFVREELVACAGLS

SPRIPLFVAVRGNVNIIMELJEGSLQIKQMGCLPEDRALYILGALQLEYL

HTRRILHGDVAKNDVLLSDGSRAALCDFGHALCLQPDGLGKSLITGYDIIPGTETHMA

PVWVKQPCDAKVDIWSGCCMMHMLNGCHPWQVFRGLCLKIASBPPEPIREIPSC

APLTAQAQIEGLRKEPVRASAMELRRVKALQEVGGLKSPWKEVKEPRPPDQA

TCHQTLPTPEPNPAKANTDGAPEQPPPLPEPEPSKAPALNLSKESGTWEPLPL

SLSDPATKGPSFPDRATLPELELQLELIELFNLSQLPFSLSEEQILSLCSIDSL

SUSDDSEKNPSKASQSRDLSGSHVSNQAEARTCSALARGRPDIPIPSYNGV

KVQIQSLNGEHLHIREFHRVKYGDITGTISQIPATFSLVTKDQGPVCYDMEVPDSG

IDLOCTLAPDGSFAMTWVRVKGQLENRP"

4157..4162

polyA_signal

/note="putative"

4174

polyA_site

/note="putative"

ORIGIN

Query Match 45.7%; Score 2099.4; DB 3; Length 4174;

Best Local Similarity 81.7%; Pred. No. 0;

Matches 2512; Conservative 0; Mismatches 526; Indels 37; Gaps 6;

QY 201 GTCTCTCTCAGGATGAGCAACAGCTGGGAGATGGCAGTGATGGAAATGGCTGCCAGG 260

DB 54 GCGCGTCCGCGGTGAGCAGCAGCGCTGGGAGATGGCGGTGATGGAGTGGCTGCCCGG 113

QY 261 TGCCCTTGCTCAGAGTGGGCGACAGAGAAATCCCCAGCCAAAGGAGAGACGCC 320

DB 114 CACTCTCGGTCAGCAGTCGGGACAGAGAGAGCTTGCCTAAAGCAAGGAGAGACACA 173

QY 321 GCCACTGGGGAAGAAACAGAGCTCCGTCTACAGCTTGAGGCCGTGGGAGAGAGCCCTGT 380

DB 174 GTCACTGGGAGAGAGAGCTGTCATCTTAAGCTTGAGGCCGTGGGAGAGAGCCCGT 233

QY 381 GTTCTGCGGAAAGTGGGAGATCCTCAATGACGTGATTACCAAGGCGACAGCCAGGAAGG 440

DB 234 GTTCTGTGGGAAGTGGGAGATCCTAAACGACGTGATACCAAGGCGACAGCCAGGACGG 293

QY 441 CTCGAGGCGAGGCGGAGCTGCGCATCTCTATCATCGCCAGGCTGAGTGTGAGAATAGCCA 500

DB 294 CTCGAGGCGAGGACACCGGCCCATCTCCATCATCGCCAGGCTGAATGTGAGAATAGCCA 353

QY 501 AGAGTTACGCCCCACCTTTTCAGAACGATTTTTCATCGTGGTCCAAACAGTACAGCCA 560

DB 354 AGAGTTACGCCCCACCTTTTCAGAGCGCATTTTTCATCGCGGGTTCACAGCAGTACAGCCA 413

QY 561 GTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATGTCTCAGAGGCGAAATGSC 620

DB 414 GTCTGAGAGTCTCGATCAAAATCCCCCAACATGTGGCCCATGCACTGAAGGCGAAATGSC 473

QY 621 CCGTGTGTGTGGAAAGGAAAGCGTCGACGAAAGCCCGCGGAAAGAAACGGAAGAAGAG 680

DB 474 CCGTGTGTGCGGAGGGGAAACGTCGCGCAAGAGCCCGGAAAGAAACGTCGGAAGAAGAG 533

QY 681 CTCAAAGTCCCTGGCTCATGACGAGTGGCCCTTTGGCCAAAACCCCTCCCGAGGACCCCTGA 740

DB 534 CTCGAGGTCACTGGGCCCGGAGGAGTGGCCCTTAGCCAAAGCCCTTGCCGAGAACCCCTGA 593

QY 741 GCAGGAGAGTGCACCATCCAGTGCAGGAGGATGAGTCTCCACTCGGCGCCCATATGT 800

DB 594 GCAAGAGAGTGTACCATCCAGTACAGGAAGATGAGTCTCCACTGAGCAACCTCTATGCG 653

QY 801 TAGAAACACCCCGCAGTGTACCAAGCCTCTGAAAGGAAACAGGCGCTTGGGCAACTCTGTTT 860

DB 654 CAGAAATGCTCCAGTTTCAACAGCCTCTGGGGGAGACAGGCGCTTGCCACCTGTGCTT 713

QY 861 TAAGCAGCTTGGCGAGGCGCTACGGCGGCTCTGGCTCGATCAGAACTCCCAAACTGAT 920

DB 714 TAAGAAACAAATGAAGCGCTTCGACCGGTACTGCTCGACAGAACTCCCAAACTGAT 773

QY 921 CAGCCCTTGCATAGTCTGAAACCAAGTGTGAAACTGACCAACCCAGGAGCGGAGGCC 980

DB 774 CAGCCCTTGCATAGTCTTAACCAAGTGTGAAACTGACCAACCCAGGAGCGGAGGCC 833

QY 981 CCGTCCCTGCGCAGCAACCCCTTCCCTATAGCAGATGCTCATCTCTTCCCAATCCA 1040

DB 834 CCGGCCCCACCGCATCAACCCCTTCCCTACAGGGAATGCCCCATCTTTTCCCAATCTA 893

QY 1041 CCCTCTCAGCCCTGGAACCTCAACCTCTCGAGT---CCTTCTGGGCAACTGGCCTG 1097

DB 894 CCCCCTGGAGCCCTGGAAACCCCTATATGCTGGACTCTGCGCTCTGGCAAACTAGCGG 953

QY 1098 TGTAGACAGCCAGAAACCCCTTGCCTGACCCA---CACCTGAGCAAACTGGCCTGTGTAGA 1154

DB 954 TGTAGCGGCCAGCGGCCCTCTGCTGGGCCACCGCATCTAAGCCAACTGGGCCCATGGAGA 1013

QY 1155 CAGTCCAAAGCCCTGCTGCGCCACACCTGGAGCCAGCTGCTGTCTGTGTTGTCCTA 1214

DB 1014 CAGTCAGAAAGCGCTGCTGCGCCACACCTGGAGTCCAGCTGCGCGTCTCGGGGTGCGCT 1073

QY 1215 TGAGAGTCTTCTGTGGAGGAATACCTAGTCGATGCTCTGCAAGCGACGCTGAGCTCAAG 1274

DB 1074 AGAAAGGTTCCTCGTGGAGGAATACCTGTGTGATGCGCTCCAAAGGAAGTGTGAGCTCAGG 1133

QY 1275 CCAGGCCACAGCTGACACCGCTGGCCAAACCTGGGCGAGCAGGGGCTCCAGATCCCG 1334

DB 1134 CCAGGCCACAGCTGCGCCAGCTGGCTAAGACATGGTCTCTCGGGAAGCGCCAGCTGCA 1193

QY 1335 GGAGGCCAGCCCAAAACTGAGGACACAGGGGTGTCTGTCTCACTGAGAAATCTCAAGCC 1394

DB 1194 GAGGCTCGGCCCCGAAACTGAGGACACAGGGGGTCTCTGTCTTACTGAGAAACTCAAGCC 1253

QY 1395 AGTGATTTATCAGTACCGAGGAAGTCCACTGTGGGCCACGACCCAGCTCCCGCTGGGCGAG 1454

DB 1254 AGTGATTTATGATATCGAGGAAGGTCCACTGGATGACACACCGCTCGGGTGGGCGAG 1313

QY 1455 AGGCTCTTTCGAGAGGTGACAGGATGGAGGACAAAGCAGAGCTGGCTTCCAGTGCCTGT 1514

||||| 1314 AGGCTCTTTCGCGAGGTCACAGAAATGAAGACAAGCAGACAGGCTTCCAGTGTGCTGT 1373
QY 1515 CAAAAGGTTCGCGCTGGAAGTATTTTCGGGACAGAGAGCTGATGGCATGTGCAGGATTGAC 1574
Db 1374 CAAAAGGTTCGAGCTCGAGGTGTTTCGGGTAGAGAACTAGTGGCTGTGCTGGTCTGAG 1433
QY 1575 CTCACCCAGAAATTTGCCCTTTGTATGGAGCTGTGAGAGAAGGCTTTGGGTCAACATCTT 1634
Db 1434 CTCGCCAGAAATTCGCTCTCTATGGAGCTGTGAGAGAAGGCCGCTGGGTGAACATCTT 1493
QY 1635 CATGGAGCTGTGGAAGTGGCTCCCTGGGCCAGCTGCTCAAGGAGCAGGCTGTCTCCC 1694
Db 1494 CATGNACTGTAGAAAGTGGCTCCTGGGTGAGCTCATAAAGCAAAATGGGCTGTCTGCC 1553
QY 1695 AGAGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTGGAGGGTCTGGAATACCTCCACTC 1754
Db 1554 AGAAGCCGAGCCCTTTACTACTCTGGGCCAGGCCCTGGAGGGCTGGAGTACTCCACAC 1613
QY 1755 ACGAAGGATTTGCATGGGGAGCTCAAAAGCTGACAACTGCTCTCTGTCACGCGATGGAG 1814
Db 1614 ACGCAGGATTTGCATGGCGATGTCAAAAGCTGACAACTGCTCTCTGTCACGATGGAAG 1673
QY 1815 CCAGCAGCCCTCTGTGACTTTGGCCATGCTGTGCTTCAACTGATGCGCTGGAAA 1874
Db 1674 CCGAGCGCCCTCTGCGACTTTGGCCAGCCCTTTGCTGCAACTGAGCGCCCTAGGGAA 1733
QY 1875 GTCTTGTCTACAGGGGACTACATCCCTGGCACAGAGACCCACATGCTCCCGAGGTGGT 1934
Db 1734 ATCTTGTCTACAGGGGACTACATCTCTGGCACGGAGACCCACATGGCACCAAGTGGT 1793
QY 1935 GCTGGCAGGAGCTCGACGCGCAAGTGGATGTCTGGAGCAGCTGCTGTATGATGCTGCA 1994
Db 1794 GATGGAAAGCCCTCGCATGCGCAAGTGGACATCTGGAGCAGCTGCTGATGATGCTCCA 1853
QY 1995 CATGCTCAACGGCTGCGACCCCTGGACTCAGTTCTTCCGAGGGCGCTCTCCCTCAAGAT 2054
Db 1854 CATGCTCAACGGCTGCGACCCCTGGACTCAGTTCTTCCGAGGGCGCTCTCCCTCAAGAT 1913
QY 2055 TGCCAGCAGGCTTCGCGCTGTGAGGGAGATCCACCCCTCTCGCGCCCTCTCAAGGCCA 2114
Db 1914 TGCCAGCAGGCTTCACCGATCAGGGAGATCCACCTCTCTCGCGCACCCCTCAAGGCCA 1973
QY 2115 GGCCATCAAGAGGGGCTGAGGAAAGGCCCATCCACCGCTGTCTGAGCGGAGCTGGG 2174
Db 1974 GGCCATCAAGAGGGGCTGAGGAAAGGCCCGTCCACCGAGCATCTGCGATGGAGCTTGG 2033
QY 2175 AGGNAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGAGCCCTTGGAGGGAGA 2234
Db 2034 GAGGAAAGTGGGCAAGGCATACAGGAAGTGGAGGTCTGAAAAGCCCTTGGAAAGGAGA 2093
QY 2235 ATATAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATTACCAACAGACCCCTCCATGC 2294
Db 2094 ATATAAGAAACAAGACCTCCACCAAGACAAGCCACCTGCGCACCAAGCCCTTACCTAC 2153
QY 2295 CCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGGCGCAGCTGAGGAGACAAAGG 2354
Db 2154 TCCGCCGAGAGA-----GAAACCCACAGCCCAAGGCCCAACACAGA 2192
QY 2355 CAGAGCCCTTAAGCTCAGACCTCTCTCCCAACAGAGCCCCACAGGCCCAACAGTCTCC 2414
Db 2193 CGGGGCTCTGAGCCTCAGCCTCTCTACCGCCAGAACCCACAGAACCGAGCAAGGGCC 2252
QY 2415 TCCCTTGACTTTGAGCAGGAGGAGTCTGGATGTGGAAACCTTACTCTGCTCTCCCT 2474
Db 2253 AGCCCTGAACCTGAGCAAGGAGGAGTCTGGCAATGGGAAACCCCTGCTCTCTCCCT 2312
QY 2475 GGAGCCAGCCCTCTGCAGAAACCCAGCTCACAGAGCGGAAAGCAACCGTCCCGGAGCA 2534
Db 2313 GGAACCCGCACTGCCAAGGCCCCAGCTTCCAGACCGGAGGGCAACCTTGGCAGAGCT 2372
QY 2535 GGAATCTGAGCAGCTGGAAATAGAAATTTATCTCTCAACAGCCTGTCCAGGCCATTTCTCT 2594
|||||

Db 2373 GGAGCTACAGCAACTGGAGATAGAACTGTTTCTCAACAGCCTGTCCAGCCGTTCTCTCT 2432
QY 2595 GGAGGACGAGGACAAATTTCTCGTGCCTCAGCATCGACAGCCTCTCTCTGTGCGATGA 2654
Db 2433 GGAGGAAACAGGAACAAATCTCTCTCTCTCAGCATCGACAGCCTCTCGCTGTGAGATGA 2492
QY 2655 CAGTGAGAAGAACCCATCAAAAGGCTCTCAAAAGCTCGGGGACACCCCTGAGCTCAGGGCT 2714
Db 2493 CAGTGAAGAAATCACTGAAGGCTCTCAGAGCTCAGGGACACCCCTGAGTTCTGGCGT 2552
QY 2715 ACATCTCTGGAGCAGCAGGCCGAGGCTCGAAGCTCTCAGCTTGGAAACATGGTGTGGCCCG 2774
Db 2553 GCATCTTTGGAAACAGCAGCTGAGGCAAGAACTGAGCTGAGCAGCGGCTGTGGCCCG 2612
QY 2775 GGCGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAGTCCAAATACAGTCTCT 2834
Db 2613 GGCGCGGCTTACTGACATCCGAGCTCAACGGGTCAGAGTCCAGATCCAGTCTCT 2672
QY 2835 TAATGGTGAACACTCTGACATCCGGGAGTTCCACCGGTTCAAGTGGGAGACATCGCCAC 2894
Db 2673 CAATGGCGAAACACTGCGATATCCGGGAATTCACCGGTCAGAGTGGGAGACATTTGCCAC 2732
QY 2895 TGGCATCAGCAGCCAGATCCAGCTGCGAGCTTCCAGCTTGGTCAACAAAGACGGCAGCC 2954
Db 2733 CGGCATCAGCAGCCAGATCCAGCCACAGCTTTTTCAGCTGTGTGACAAAGATGGACGCC 2792
QY 2955 TGTTCGTACGACATGGAGGTGCCAGACTCGGGCATCGACTGAGTGCACACTGGGCC 3014
Db 2793 TGTTCGTATGACATGAGGTGCCAGACTCGGGCATCGACCTGCGAGTGCACCTGGGCC 2852
QY 3015 TGATGGCAGCTTCGCTTGGAGCTGGAGGTTCAAGCATGGCCAGCTGAGGAAACAGGCCCTA 3074
Db 2853 TGATGGCAGCTTTTCTTGGACCTGGAGGTTCAAGCATGGTTCAGCTGGAGAACCGACCTA 2912
QY 3075 ACCCTGCTCCACCGCGCTCCACACTG---CCGGAAGCAGCCTTCTCTGCTCGGTGC 3131
Db 2913 GTCTGTCTTCCACCAACAGCTCCACTCTGCGCACGGGCTTCTTTGATCTCAGTGC 2972
QY 3132 ACATGCTGCTCCGTAACACAGGCTCAGCGCTTCCAGGGGATTCGAGCCCGCCCGCT 3191
Db 2973 TCGTGTGTCTGAGATGCTGG---ATGAGTCTGCACAGGACCTGCCAGTCCCTGGTT 3029
QY 3192 CACAGTGGGAACCAAGGCTTCGAGCAGCAAGTGGGGGCA----AGCAGAAATGCTTCCC 3247
Db 3030 CAGCATGTAGACCAAGGCTTGTGAGCAGCAGCAGCAAGGGAACAGAGTAAGACACCTCCC 3089
QY 3248 AGGATTTACACCTG 3262
Db 3090 AGGACTTGGCTCTCTG 3104

RESULT 4

AY413578 LOCUS 2765 bp DNA linear GSS 17-DEC-2003
DEFINITION Pan troglodytes MAP3K14 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

ACCESSION AY413578

VERSION AY413578.1 GI:39769540

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes

REFERENCE 1 (bases 1 to 2765)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarial,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios

JOURNAL Science 302 (5852), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 2765)

AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.	Db	721 TCCATTCCACCCCTCTCCAGCCCTTGAAAACCTCACCCCTCTGGAGTCTCTTCTCTGGGCAAC 780
TITLE	Direct Submission	Qy	1091 TGGCTCTGTAGACAGCAGCAAAACCCCTTGCCTGACCCACACCTGAGCAAACTGSCCTGTG 1150
JOURNAL		Db	781 TGGCTGTGTNNCAGNCNCAACCCCTTGCCTGACCCACACCTGAGCAAACTGSCCTGTG 840
COMMENT	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment.	Qy	1151 TAGACAGTCCAAAGCCCTCTGCTGGCCCAACACTGGAGCCAGCTGCTCTCTCTGCTGTG 1210
FEATURES	Location/Qualifiers	Db	841 TNNNCAGTCCAAAGCCCTCTGCTGGCCCAACACTGGAGCCAGCTGCTCTCTCTGCTGTG 900
source	1..2765 /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598" <1..>2765 /gene="MAP3K14" /locus_tag="HCM4937"	Qy	1211 CCATGAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGAGCAGCTGAGCT 1270
gene		Db	901 CCCATGAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGAGCAGCTGAGCT 960
ORIGIN		Qy	1271 CAAGCCAGGCCACAGAGCTGACAGCTGGCCAAAGCTCTGGGAGCAAGGGCTCCAGAT 1330
		Db	961 CAGGCCAGGCCACAGAGCTGACAGCTGGCCCAAGACTCTGGGAGCAAGGGGCTCCAGAT 1020
		Qy	1331 CCCGGAGCCAGCCCAAAACTGAGGACAAAGAGGGTGTCTCTGCTCCTGAGAAACTCA 1390
		Db	1021 CCCGGAGCCAGCCCAAAACTGAGGACAAAGAGGGTGTCTCTGCTCANNNNAAAATCA 1080
		Qy	1391 AGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCCACGACACAGCTCCGCTGG 1450
		Db	1081 AGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCCACGACACAGCTCCGCTGG 1140
		Qy	1451 GCAGAGGCTCTTCGAGAGAGTGACAGGATGGAGGACAAAGACAGACTGGCTTCCAGTGG 1510
		Db	1141 GCAGAGGCTCTTCGAGAGAGTGACAGGATGGAGGACAAAGAGCTGGCTTCCAGTGG 1200
		Qy	1511 CTGTCAAAAAGTGGCTGGAAGTATTTTCGGGAGAGAGGAGCTGATGGCATGTCCAGGAT 1570
		Db	1201 CTGTCAAAAAGTGGCTGGAANNATTTTCGGGCAANNAGAGCTGATGGCATGTCCAGGAT 1260
		Qy	1571 TGACCTCACCCAGAAATTTGTCCTTTGTATGGAGCTGTGAGAGAGGCGCTTGGGTCAACA 1630
		Db	1261 TGAANNNNCAGAAATTTGTCCTTTGTATGGAGCTGTGAGAGAGGCGCTTGGGTCAACA 1320
		Qy	1631 TCTTCATGGAGCTGTGGAAGTGGCTCCCTGGGCAGCTGGTCAAGAGCAGGCGCTGTC 1690
		Db	1321 TCTTCATGGAGCTGTGGAAGTGGCTCCCTGGGCAGCTGGTCAAGAGCAGGCGCTGTC 1380
		Qy	1691 TCCAGAGACCGGGCCCTGTACTACTGGGCCAGGCGCTTGAGAGGTCTGGAATACCTCC 1750
		Db	1381 TCCAGAGACCGGGCCCTGTACTACTGGGCCAGGCGCTTGAGAGGTCTGGAATACCTCC 1440
		Qy	1751 ACTCACGAAGGATTTGTCATGGGAGCGTCAAAAGCTGACAACCGTGTCTCTGTCCAGCGATG 1810
		Db	1441 ACTCACGAAGGATTCNGCATGGGAGCTCAAGAGNNNNNNNNNNNNNNNNNNNNNNNN 1500
		Qy	1811 GGAGCCACGACGCTCTGTGACTTTTGGCCATGCTGTGTCTTCAACCTGATGGCCTGG 1870
		Db	1501 NNN 1560
		Qy	1871 GAAAGTCTTGTGTCACAGGGAGTACATCCCTGGCACAGAGACCCACATGGCTCCGGAGG 1930
		Db	1561 GNANGTCTTGTCTNNNNNGGATTTACATCCCTGGCACAGAGACCCACATGGCTCCGGAGG 1620
		Qy	1931 TGGTGCTGGGAGGAGCTGCGACCAAGGTGGAATGCTTGAGAGAGCTGCTGTATGATGC 1990
		Db	1621 TGGTGCTGGGAGGCGCTGCGACGCAAGGTGGACGCTCTGGAGCAGCTGCTGTATGATGC 1680
		Qy	1991 TGCAATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCGCTCTGCCTCA 2050
		Db	1681 TGCAATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCGCTCTGCCTCA 1740
		Qy	2051 AGATTGCCAGCGAGCTCCGCTGTGAGGGAGATCCACCCCTCTCTCGCCCCCTCTCACAG 2110
		Db	1741 AGNN 1800
		Qy	2111 CCCAGGCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGACGGGAGC 2170
		Db	1801 NNN 1860

```
Qy 2171 TGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCTTGGAGGG 2230
Db 1861 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 2231 GAGAAATAAGAACCAAGACATCCACCGCAATCAAGCAATATACCAAGAGCCCTCC 2290
Db 1921 GAGAAATAAGAACCAAGACATCCACCGCAATCAAGCAATATATACCAAGAGCCCTCC 1980
Qy 2291 ATGCCACCGCAGAGAGCTTTTGGCAAGGGCCCCAGGGCCCCCGGCGCAGCTGAGGAGACAA 2350
Db 1981 ATGCCACCGCAGAGAGCTTTGGCAAGGGCCCCAGGGCCCCCGGCGCAGCTGAGGAGACAA 2040
Qy 2351 CAGCAGAGCCCTTAAGCTTCAGCTCTCTCTCTCCACAGAGCCCCCGGCGCAGCTGAGGAGACAA 2410
Db 2041 CAGCAGAGCCCTTAAGCTTCAGCTCTCTCTCTCCACAGAGCCCCCGGCGCAGCTGAGGAGACAA 2100
Qy 2411 CTCTCTCTCTGACCTTCAGCAAGAGAGAGCTGGGATGTGGGATGTGGGATGTGGGATGTGGGAT 2470
Db 2101 NTCTCTCTCTGACCTTCAGCAAGAGAGAGCTGGGATGTGGGATGTGGGATGTGGGATGTGGGAT 2160
Qy 2471 CCTGGAGCCAGCCCTTCCAGAAACCCAGCTCACAGAGCGGAAAGCAACGTCGCCGG 2530
Db 2161 CCTGGAGCCAGCCCTTCCAGAAACCCAGCTCACAGAGCGGAAAGCAACGTCGCCGG 2220
Qy 2531 AGCAGGAAGCTGACAGAGCTGGAATAGAAATTTCTCAACAGAGCTGTCCAGGCAATTT 2590
Db 2221 AGCAGGAAGCTGACAGAGCTGGAATAGAAATTTCTCAACAGAGCTGTCCAGGCAATTT 2280
Qy 2591 CTCTGGAGGAGCAGGAGCAAAATTTCTCTGTCCTCAGCATCGACAGCTCTCCCTGTGCG 2650
Db 2281 CTCTGGAGGAGCAGGAGCAAAATTTCTCTGTCCTCAGCATCGACAGCTCTCCCTGTGCG 2340
Qy 2651 ATGACAGTGAGAGAACCCATCAAGGCTCTCAAGGCTCGCGGACACCTGAGCTCAG 2710
Db 2341 ATGACAGTGAGAGAACCCATCAAGGCTCTCAAGGCTCGCGGACACCTGAGCTCAG 2400
Qy 2711 GCGTACACTCTGGAGCAGCCAGCGGCTCGAAGCTCCAGCTGGAACATGCTGTGCTGG 2770
Db 2401 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 2771 CCGGGGGGGCCACCGACACCCCAAGCTATTTCATGTGTGAAGTCAAAATACAGT 2830
Db 2461 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 2831 CTCTTAATGTGACACCTGCACATCCGGGAGTTCCACCGGTCCAAAGTGGGAGACATCG 2890
Db 2521 CTCTTAATGTGACACCTGCACATCCGGGAGTTCCACCGGTCCAAAGTGGGAGACATCG 2580
Qy 2891 CCACTGGCATCAGCAGCAGATCCAGCTGCAGCTTTCAGCTTGGTCAACAAAGACGGGC 2950
Db 2581 CCACTGGCATCAGCAGCAGATCCAGCTGCAGCTTTCAGCTTGGTCAACAAAGACGGGC 2640
Qy 2951 AGCTGTTCGTACGACATGAGAGTGCAGACTCGGGATCGACTGAGTGCACATGG 3010
Db 2641 AGCTGTTCGTACGACATGAGAGTGCAGACTCGGGATCGACTGAGTGCACATGG 2700
Qy 3011 CCCCTGATGCGAGCTTCGCTGGAGCTGGGGTCAAGCATGGCCAGCTGAGGAGACAGGC 3070
Db 2701 CCCCTGATGCGAGCTTCGCTGGAGCTGGGGTCAAGCATGGCCAGCTGAGGAGATAGGC 2760
Qy 3071 CCTAA 3075
Db 2761 CCTAA 2765
```

```
RESULT 5
AV413579
LOCUS Mus musculus MAP3K14 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION AV413579
ACCESSION AV413579
VERSION AV413579.1 GI:39769541
```

```
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 2829)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2829)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..2829
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2829
/genes="MAP3K14"
/locus_tag="HCM4937"
gene
ORIGIN
Query Match 42.1%; Score 1942; DB 9; Length 2829;
Best Local Similarity 79.0%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 571; Indels 27; Gaps 3;
```

```
Qy 232 ATGCGAGTGTGTAATGGCTGCTCCAGAGTGGCCCTGGCTCAGCAGTGGGGCAGCAGAAG 291
Db 1 ATGCGCGTGTGTAATGGCTGCTCCAGAGTGGCCCTGGCTCAGCAGTGGGGCAGCAGAAG 60
Qy 292 GAATCTCCCAAGCAAGGAGAGAGCCGCCACTGCGGGAAGAAACAGAGTCCGTCTAC 351
Db 61 GAGCTTGCCAAAGCAAGGAGAGAGACACAGTCACTGGGGAAGAGCAGAGTGCATCTTC 120
Qy 352 AAGCTTGAGCCCTGGAGAGAGCCCTGCTTCTCGGGAAGTGGGAGATCCTCAATGAC 411
Db 121 AAGCTTGAGCCCTGGAGAGAGCCCTGCTTCTCGGGAAGTGGGAGATCCTCAATGAC 180
Qy 412 GTGATTACCAAGGCGCAGCAAGGAGGCTCCGAGGCGAGGCGCAGCTCTCTATC 471
Db 181 GTGATCAACAAGGCGCAGCAAGGAGGCTCTGAGGAGGAGCACCGGCCATCTCCATC 240
Qy 472 ATCGCCAGAGCTGAGTGTGAGAAATAGCAAGAGTTACGCCCCACCTTTTCAGAACGCAAT 531
Db 241 ATCGCCAGAGCTGAAATGTGAGAAATAGCAAGAGTTACGCCCCACCTTTTCAGAACGCAAT 300
Qy 532 TTCTATCGCTGGTCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAT 591
Db 301 TTCTATCGCTGGTCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAT 360
Qy 592 GTGGCCCATCTACAGAGGGCAAAATGGCCCGCTGTGTGTTGGAGGGAAGCGTCGCGAGC 651
Db 361 GTGGCCCATCTCAACTGAGGCAAAATGGCCCGCTGTGTGTTGGAGGGAAGCGTCGCGAGC 420
Qy 652 AAAGCCCGAAGAAACGGAAGAGAGAGTCAAGTCCCTGGCTCATGACGAGTGGCC 711
Db 421 AAAGCCCGAAGAAACGGAAGAGAGAGTCAAGTCCCTGGCTCATGACGAGTGGCC 480
Qy 712 TTGGCCAAACCCCTCCCGCAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGCAGGAG 771
Db 481 TTAGCCAAAGCCCTCGCCAGAACCCCTGAGCAGGAGAGCTGTATCCATCCCGATACAGGAA 540
```


QY 772 GATGAGTCTCCACTCGGCGGCCCATATGTTAGAAACACCCCGCAGTTCCACAGCCTCTG 831
Db 541 GATGAGTCTCCACTAGGCAACCTCTATGCCAGAAATGTTCTCCAGTTCCACAGCCTCTG 600
QY 832 AAGAAACCGAGGCTTGGGCAACTCTGTTTTAAGCAGCTTGGGAGGCGCTACGGCGGCT 891
Db 601 GGGGACCGAGGCTTGGCCACCTGTGCTTTAAGAACAAAGATGAAGGCTTCGACCGGTA 660
QY 892 CTGCGCTCGATCAGAACTCCACAAACTGATCAGCCCTTGGAAATGTTGAAACACGCTGTGG 951
Db 661 CTGCGCTCGACCAAACTCCACAAACTGATCAGCCCTTGTCAATGTTAAACCAACGCTGTGG 720
QY 952 AAACCTCACACCCCGAGGACGAGGCGCCCTGCGCCCTGCCACGACCCCTTCCCTTAT 1011
Db 721 AAACCTCACACCCCGAGGACGAGGCGCCCGGCCCCCAGCTCACCCCTTCCCTTAC 780
QY 1012 AGCAGACTCGCTATCCCTTCCCATTTCCACCTCTCCAGCCCTTGAAACCTCACCCCTCTG 1071
Db 781 AGCGGAATGCCCATCTCTTCCATTCTACCCCTTGAGGCCCTTGAAACCTCTATATGCTG 840
QY 1072 GAGT---CCTTCTGGGCAACTGGCCTGTGTAGACAGCAAGAACCTTGGCTGACCCCA 1128
Db 841 GACTCTGCGCTCTGGACAACTAGCCGGTGTGAGCGGCCAGCGGCTCTGCTGGCCCA 900
QY 1129 ---CACCTGAGCAAACTGGCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTG 1185
Db 901 CCGCATCTAAGCCAACTGGCCCATGGAGACAGTCAAGAGCCGCTGCTGGCCCAACCTG 960
QY 1186 GAGCCAGCTGCTGTCTGCTGGTCCCATGAGAAAGTTTCTGTGGAGAAATACCTAGTG 1245
Db 961 GAGTCCAGCTGCGCGTCTCGGGTGCCTTAGAAAGGTTTCCCGTGGAGGAATACCTGGTG 1020
QY 1246 CATGCTCTCAGAGCAGCTGAGCTCAAGCCAGGCCCAAGCCTGACAGCCTTGGCCAG 1305
Db 1021 CATGGCTCTCAGGAAGTGTAGCTCAGGCCAGGCCCAAGCCTTGGCCAGCCTTGGCTAAG 1080
QY 1306 ACCTGGGACAGCGGGCTTCAGATCCCGGGAGGCCAGCCCCCAAACTGAGGACAAACGAG 1365
Db 1081 ACATGGTCTCTGGGAAGCGCCAGCTGACAGGCTCGGCCCGAACTGAGGACAAACGAG 1140
QY 1366 GGTGTCTGCTCACTGAGAACTCAAGCAGTGGATTATGAGTACGGAAGAAATCCAC 1425
Db 1141 GGGGTCTCTTACTGAGAACTCAAGCCAGTGGATTATGAGTACGGAAGAGGTCCAC 1200
QY 1426 TGGGCCACACACAGCTCGGCTGGGCGAGAGGCTCCTTCGGAGAGGTGCACAGATGGAG 1485
Db 1201 TGGATGACACACAGCTCGGGTGGCGAGAGGCTCCTTCGGCGAGGTCCACAGAAATGAAG 1260
QY 1486 GACAGCAGACTGGCTTCCAGTGCCTGTCAAAAAGGTGGCGCTGGAAGTATTTTCGGCA 1545
Db 1261 GACAGCAGACAGGCTTCCAGTGTGCTGTCAAAAAGGTACGACTCGAGGTGTTTCGGTA 1320
QY 1546 GAGGAGCTATGCGATGTCAGGATTGACTCACCCAGAAATGTGCTCCCTTTGTATGAGCT 1605
Db 1321 GAGGAACTAGTGGCTGTGCTGGTCTGAGCTCGCCAGAAATCGTCCCTCTCTATGAGCT 1380
QY 1606 GTGAGAGAGGCGCTTGGGTCAACATCTTCATGAGAGCTGCTGGAAGGTGGCTCCCTGGGC 1665
Db 1381 GTGAGAGAGGCGCTGGGTGAACATCTTCATGGAAGTCTGTAAGGTTGGCTCGCTGGGT 1440
QY 1666 CAGCTGGTCAAGGACAGGCTGTCTCCAGAGGACCGGGCCCTGTACTTACCTGGGCCAG 1725
Db 1441 CAGCTCATAAAGCAAAATGGGCTGTCTGCCAGAAAGCCGAGCCCTTTACTACCTGGGCCAG 1500
QY 1726 GCCCTGGAGGCTCTGGAATACCTCACTCAAGAGGATCTGCATGGGGAAGTCAAAGCT 1785
Db 1501 GCCCTGGAGGCGCTGGAGTACCTCCACACACGACGAGTCTGCATGGCGATGTCAAAGCT 1560
QY 1786 GACAACTGCTCCTCTCAGCGATGGGAGCCACGACGCTCTGTGATCTTTGGCCATGCT 1845
Db 1561 GACAACTGCTCCTCTCAGTGTGGAAGCCGAGGCGGCTCTGCGCATTTGGCCACGCCC 1620

QY 1846 GTGTCCTCTCAACCTGATGGGCTGGGAAAGTCTTGTCTCAGAGGAGCTACATCCTGGC 1905
Db 1621 TTGTGCTCTGCAACCTGACGGCCTAGGGAAATCTCTTGTCTCAGAGGAGCTACATCCTGGC 1680
QY 1906 ACAGAGACCCACATCGCTCCGAGGTGTCTGGGACGAGCTGCGCAAGGTGGAT 1965
Db 1681 AGGGAGACCCACATGCGCACCAAGTGTGTATGGGAAAGCCCTGCGATGCGCAAGGTGGAC 1740
QY 1966 GTCTGGAGAGCTGTGTATGATGTCACATGTCACACGCTGCCACCTGGAGCTCAG 2025
Db 1741 ATCTGGAGCAGCTGTGATGATGTCCTCAGATGTCCTCAGCGGTGCGCACCTGGAGCTCAG 1800
QY 2026 TTCTTCCGAGGCGCTCTGCTCAAGATTGCCAGCAGCCTCCGCTGTGAGGGAGATC 2085
Db 1801 TACTTCCGAGGCGCTTGTCTCAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1860
QY 2086 CCACCTCTCTGCGCCCTCTCTCACAGCCAGGCGCATCAAGAGGGGCTGAGGAAAGAGCCC 2145
Db 1861 NNN 1920
QY 2146 ATCCACCGCTGTCTGACGCGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTG 2205
Db 1921 NNTG 1980
QY 2206 GGAGTCTGAAAGAGCCCTTGGAGGGGGAATATAAGAAACCAAGACATCCACCGCAAAT 2265
Db 1981 GGAGTCTGAAAGCCCTTGGAAAGGAGAAATATAAGAAACCAAGACCTCCACCCAGAC 2040
QY 2266 CAAGCAATTAACACAGACCTCTCATGCCAGCGCAGAGAGCTTTGCCCAAGGGCCCCA 2325
Db 2041 CAAGCACTCTGCCACAGACCTTACTTCTCGCGCAGAGA----- 2081
QY 2326 GGGCCCCGGCAGCTGAGGAGACACAGGAGAGGCGCTTAAGCTCAGCCTCTCTCCCA 2385
Db 2082 --GAACCCACAGCCCAAGGCCAACACAGAGGGGCTCTGAGCCTCAGCCTCTCTACCG 2139
QY 2386 CCAGAGCCCCCAGAGCCAAACAAAGTCTCCTCCCTTGTACTTTGAGCAAGAGGAGTCTGG 2445
Db 2140 CCAGAACCCACAGAACCGAGCAAGGCGCAGCTTGAACCTGAGCAAGAGGAGTCTGGC 2199
QY 2446 ATGTGGAAACCTTACTCTCTCTCTGAGCGCAGCCCTGCGCAGAAACCCCAAGTCA 2505
Db 2200 ACATGGAAACCTCTGCTCTCTCTGAGCCAGCCACTGCGCAAGGCCCCCAGCTTC 2259
QY 2506 CCAGAGGGAAAGCAACGCTCCGAGCAGGAACTGAGCAGCTGGAATATGATTTTC 2565
Db 2260 CCAGAGCGGGCAACCTTCCAGAGCTGAGCTTACAGCAACTGGAGATGAACTGTGT 2319
QY 2566 CTCAACAGCTGTCTCCAGCCATTTTCTGAGGAGCAGAGCAAAATTTCTCTGTCGCTC 2625
Db 2320 CTCAACAGCTGTCTCCAGCCGTTCTCTGAGGAAACAGGAAACAAATCTCTCTGCTC 2379
QY 2626 AGCATCGACAGCTCTCTCTCTGAGTGAACAGTGAAGAAACCCATCAAGGCTCTCAA 2685
Db 2380 AGCATCGACAGCTCTCTCTGATGACAGTGAAGAAATCCATCGAAGCCTCTCAG 2439
QY 2686 AGCTCGGGGACACCTGAGCTCAGGCTACACTCTGAGCAGCCAGCGCGAGGCTCGA 2745
Db 2440 AGCTCAAGGAGACACCTGAGTCTGGGCTGCACTCTTTGGAACACGCAAGCTGAGGCAAGA 2499
QY 2746 AGCTCAGCTGGAAACATGGTGTCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTC 2805
Db 2500 ACCTGAGCTGACAGCGGCTGCGCCCGGGGGCGGCTACTGTGACATCCCGAGCTACTTC 2559
QY 2806 AATGCTGAAGTCCAAATACATGCTCTTATGCTGAACACCTGCGCATCCCGGAGTTC 2865
Db 2560 AACGGGTCAAGTCCAGATCCAGTCTCTCAATGGCGAAACACCTGCGATATCCGGGAAATTC 2619
QY 2866 CACCGGTCAAAGTGGGAGACATCGCACTGGCATCAGACCCAGATCCAGCTGCGAGCC 2925
Db 2620 CACCGGTCAAAGTGGGAGACATTCGCCATCAGAGCCAGATCCCGAGCCACAGCT 2679
QY 2926 TTCAGCTTGGTCAACCAAGAGCGGCGAGCCTGTTCTGCTACGACATGGAGGTGCCAGACTCG 2985


```
Db      2680 TTACGCTGGTGACCAAGATGGACAGCCTGTTGCTATGACATGGAGGTGCCAGACTCG 2739
Qy      2986 GGCATCGACCTGCAGTGCACACTGGCCCTGATGGCAGCTTTCGCTGGAGCTGGAGGGTC 3045
Db      2740 GGCATCGACCTGCAGTGCACCTGGCCCTGATGGCAGCTTGTGTTGACCTGGAGGGTC 2799
Qy      3046 AAGCATGCCAGCTGGAGAAACAGGCCCTA 3074
Db      2800 AAGCATGCTCAGCTGGAGAACCGACCTA 2828

RESULT 6
LOCUS   CR612342                1640 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSODK011Y14 of HeLa cells Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR612342
VERSION   CR612342.1 GI:50493149
KEYWORDS  HTC; CNSLT cDNA.
SOURCE   Homo sapiens
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1. (bases 1 to 1640)
AUTHORS   Li.W.B., Gruber.C., Jesse.J. and Polayes.D.
JOURNAL   Full-length cDNA libraries and normalization
          Unpublished
REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
          http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
          Faraday Avenue
          Genoscope.
REFERENCE 2. (bases 1 to 1640)
AUTHORS   Genoscope.
JOURNAL   Direct Submission
          Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
          BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr
          - Web : www.genoscope.cns.fr)
COMMENT   1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
          end enriched, double-strand cDNA was digested with Not I and cloned
          into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
          was normalized. Library was constructed by Life Technologies, a
          division of invitrogen.
          Location/Qualifiers
FEATURES             source
     source           1..1640
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CSODK011Y14"
                     /tissue type="Hela cells Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"

ORIGIN
Query Match 33.1%; Score 1521.2; DB 3; Length 1640;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1615; Conservative 0; Mismatches 23; Indels 10; Gaps 7;

Qy      2922 AGCTTCAGCTTGTGCACCAAGACGGCAGCCTGTTGCTACGACATGGAGGTGCCAGA 2981
Db      1 AGCTTCAGCTTGTGCACCAAGACGGCAGCCTGTTGCTACGACATGGAGGTGCCAGA 60
Qy      2982 CTCGGGCATCGACCTGCAGTGCACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGGAG 3041
Db      61 CTCGGGCATCGACCTGCAGTGCACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGGAG 120
Qy      3042 GGTCAAGCATGGCAGCTGGAGAAACAGGCCCTTAACCTGTCCTCACCGCCGCTCCACA 3101
Db      121 GGTCAAGCATGGCAGCTGGAGAAACAGGCCCTTAACCTGTCCTCACCGCCGCTCCACA 180
Qy      3102 CTCGGGAACAGCCTTCCTGCTGGTGACGATGCTGCCCTGAACACAGAGCTCAGC 3161
Db      181 CTCGGG-AAGAGCCTTCCTGCTGGTGACGATGCTGCCCTG-AAACAGAGCTCAGC 238
Qy      3162 CGTTCCAGGGGATTGCCAGCCCCCGGCTCACAGTGGGAACCGAGGGCTCGCAGCAGCA 3221
```

```
Db      239 CGTTCCAGGGGATCTGCCAG-CCCCCGGCTCAGCAGTGGGACCGAGGGCTTCGACAGCA 297
Qy      3222 AGTGGGGGGAAGCAGAAATGCTCCAGAGATTTCACACTCAGCCTGCTCCGCCACCCCTGCT 3281
Db      298 AGTGGGGGGAAGCAGAAATGCTCCAGAGATTTCACACTCAGCCTGCTCCGCCACCCCTGCT 357
Qy      3282 GAAAAAACA-TCCGCCACGTGAAGAGACAGAAAGAGGATGGCAGGAGTTACTCTGGGAAA 3340
Db      358 GAGAAAAACAATCCGCCACGTGAAGAGACAG-AGGAGGATGGCAGGAGTTACTCTGGGAAA 416
Qy      3341 CAAAAACAGGATCTTTTCTGCCCCCTGCTCAGTGCAGTGGCCCTGAGCCCCCTTGGATCA 3400
Db      417 CAAAACAGG-ATCTTCTCTGCCCCCTGCTCAGTGCAGTGGCCCTGAGCCCCCTTGGATCA 472
Qy      3401 GTGACCAATTTGTCGACAGAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGAGGGGTGGGC 3460
Db      473 GTGACCAATTTGTCGACAGAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGAGGGGTGGGC 532
Qy      3461 GAGCCCTTCGGCCCCCTCACCCCTCCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCCA 3520
Db      533 GAGCCCTTCGGCCCCCTCACCCCTCCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCCA 592
Qy      3521 AACTCAGGTTTCAGTGCAGAACCCAGGTGATGCCGCCCTGCTAGGTAAAGGGGCC 3580
Db      593 AACTCAGGTTTCAGTGCAGAACCCAGGTGATGCCGCCCTGCTAGGTAAAGGGGCC 652
Qy      3581 CTCTAAACCCCTTGGCTGGCCCTCACCTGGCCAGCTCACCCCTTTTGGGTAGGGGAAA 3640
Db      653 CTCTAAACCCCTTGGCTGGCCCTCACCTGGCCAGCTCACCCCTTTTGGGTAGGGGAAA 712
Qy      3641 GAATGCTGACCCCTGGGAAGGCTCCCTGGTAGAATAACACACACTTTTCAGGTTGTTGCA 3700
Db      713 GAATGCTGACCCCTGGGAAGGCTCCCTGGTAGAATAACACACACTTTTCAGGTTGTTGCA 772
Qy      3701 ACACAGTCTCTGAGTTGACCTCTGGTTACGCAAGGACCAAGAGGTGTGTAGTGAAG 3760
Db      773 ACACAGTCTCTGAGTTGACCTCTGGTTACGCAAGGACCAAGAGGTGTGTAGTGAAG 832
Qy      3761 TGGTTTCAGTCCCCACAGCATGTCCTTTCGCTGGCTGACCTCTTCCCCAGAGCA 3820
Db      833 TGGTTTCAGTCCCCACAGCATGTCCTTTCGCTGGCTGACCTCTTCCCCAGAGCA 892
Qy      3821 GCAGGCCCCGAGCCCTTCAGGGCCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCT 3880
Db      893 GCAGGCCCCGAGCCCTTCAGGGCCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCT 952
Qy      3881 CATCTGTAAAGGTGAAGGGTGTATGAGGATATGCTGACAGGAAACAGTCTGTGTGATGAC 3940
Db      953 CATCTGTAAAGGTGAAGGGTGTATGAGGATATGCTGACAGGAAACAGTCTGTGTGATGAC 1012
Qy      3941 ATGATCAGTCTAAGGAAACAGCAGAGAGAGCG-TCCGGCGCCCGAGCCCACTATCA 3999
Db      1013 ATGATCAGTCTAAGGAAACAGCAGAGAGAGAGCGTCCGGCGCCCGAGCCCACTATCA 1072
Qy      4000 GTGTCCAGCGTGTGTTTCCCGAGAGCAGCTCAGCATACACTGACACTCACCTGCC 4059
Db      1073 GTGTCCAGCGTGTGTTTCCCGAGAGCAGCTCAGCATACACTGACACTCACCTGCC 1132
Qy      4060 CTGCCCCCTGGCCAGAGGGTACTGCCGACGGCACTTTGCACTCTGATGACCTCAAGCACT 4119
Db      1133 CTGCCCCCTGGCCAGAGGGTACTGCCGACGGCACTTTGCACTCTGATGACCTCAAGCACT 1192
Qy      4120 TTGATGCTGCCCTCTGGCAGGGCAGGGCAGGTGACACTGTAGGAGCATAGCAGC 4179
Db      1193 TTGATGCTGCCCTCTGGCAGGGCAGGGCAGGTGACACTGTAGGAGCATAGCAGC 1252
Qy      4180 CAGGAGATGGGTGAAGGGACACAGTCTTGAGCTGTCCATGATGTGACTCTCTCAAC 4239
Db      1253 CAGGAGATGGGTGAAGGGACACAGTCTTGAGCTGTCCATGATGTGACTCTCTCAAC 1312
Qy      4240 CTCTTCCAGATTTCTCTAAGAATAAGCCCCCTTCCCAATGGCCAGCTTAGCCTCTTC 4299
```


LOCUS AL556130 1133 bp mRNA linear EST 30-MAR-2004
DEFINITION AL556130 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DK011Y14 5-PRIME, mRNA sequence.

ACCESSION AL556130
VERSION AL556130.3 GI:45860849
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1133)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31277934.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8426.r

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DK011Y14&q=8426.r.

FEATURES
Location/Qualifiers
1..1133
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011Y14"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6
vector. Library was normalized."

ORIGIN
Query Match 19.9%; Score 916; DB 1; Length 1133;
Best Local Similarity 95.3%; Pred. No. 1.9e-219;
Matches 1058; Conservative 4; Mismatches 34; Indels 14; Gaps 11;

Qy 2922 AGCTTCAGCTTGTCACCAAGACGGCAGCGCTTTCGTCACGATGGAGTGCACGA 2981
Db 1 AGCTTCAGCTTGTCACCAAGACGGCAGCGCTTTCGTCACGATGGAGTGCACGA 60
Qy 2982 CTCGGGCATCGACCTCGAGTCACACTGGCCCTCGATGGCAGCTTCGGCTGGAGTGGAG 3041
Db 61 CTCGGGCATCGACCTCGAGTCACACTGGCCCTCGATGGCAGCTTCGGCTGGAGTGGAG 120
Qy 3042 GGTCAAGCATGGCAGCTGGAGAACAGAGCCCTAACCTTCGCTCCACCGCGGTCCACA 3101
Db 121 GGTCAAGCATGGCAGCTGGAGAACAGAGCCCTAACCTTCGCTCCACCGCGGTCCACA 180
Qy 3102 CTGCGGGAAGCAGCTTCCTCGCTCGGTGCAGCATGCTGCCCTGAACACAGAGCTCAGC 3161
Db 181 CTGCGGGAAGCAGCTTCCTCGCTCGGTGCAGCATGCTGCCCTGAACACAGAGCTCAGC 238
Qy 3162 CGTTCCAGGGGATTGCCAGCGCCCGCTCACAGTGGGAACACAGGGCTCGCAGCAGCA 3221
Db 239 CGTTCCAGGGGATTGCCAG -CCCCCGCTCAGCAGTGGGACAGGGCTCGCAGCAGCA 297
Qy 3222 AGGTGGGGGCAAGCAGATGCTCCCGAGATTTCACACCTGAGCCCTGCCCCACCTGCT 3281
Db 298 AGGTGGGGGCAAGCAGATGCTCCCGAGATTTCACACCTGAGCCCTGCCCCACCTGCT 357
Qy 3282 GAAAAAACA-TCCGCGCAGTGAAGAGACAGAGGAGGATGCGAGGATTACTCGGGAAA 3340
Db 358 GAAAAAACA-TCCGCGCAGTGAAGAGACAG -AGGAGGATGGCAGGATTACTCGGGAAA 416

Qy 3341 CAAAACAGGGATCTTTTTCGCCCTCTGCTCCAGTCGAGTTGGCTGACCCCGCTTGGATCA 3400
Db 417 CAAAACAGG----ATCTTCTCTGCTGCCCTCTGCTCCAGTCGAGTTGGCTGACCCCGCTTGGATCA 472
Qy 3401 GTGACCATTTGTTGGCAGACAGGGGAGAGCAGCTTCCAGCCCTGGGTCAAGAGGGGTGGGC 3460
Db 473 GTGACCATTTGTTGGCAGACAGGGGAGAGCAGCTTCCAGCCCTGGGTCAAGAGGGGTGGGC 532
Qy 3461 GAGCCCTTCGGCCCTCACCCCTCCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCCA 3520
Db 533 GAGCCCTTCGGCCCTCACCCCTCCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCCA 592
Qy 3521 AACTCAGGTTCAAGTCAGAACACAGGTGAGCAGGTATGCCGCCCTGATGTTAAAGGGGGCC 3580
Db 593 AACTCAGGTTCAAGTCAGAACACAGGTGAGCAGGTATGCCGCCCTGATGTTAAAGGGGGCC 651
Qy 3581 CTCTAAACCCCTTCGCTGGCTCACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAA 3640
Db 652 CTCTAAACCCCTTCGCTGGCTCACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAA 711
Qy 3641 GAATGCTGACCCCTGGGAAGGCTCCCT-GGTAGATATACACACACTTTTTCAGGTTGTGC 3699
Db 712 GAATGCTGACCCCTGGGAAGGCTCCCTGGGTAGATATACACACACTTTTTCAGGTTGTGC 771
Qy 3700 AACACAGGTCCTGAGTTGACCTCTGTTTCAGCCCAAGGACCAAGAGGTGTGAAGTCAA 3759
Db 772 AACACAGGTCCTGAGTTGACCTCTGTTTCAGCCCAAGGACCAAGAGGTGTGAAGTCAA 831
Qy 3760 GTGGTTCTCAGTCCCGCAGACATGTGCCCTTTGCTGTGGTACCACTCTTCCCGCAGAC 3819
Db 832 GTGGTTCTCAGTCCCGCAGACATGTGCCCTTTGCTGTGGTACCACTCTTCCCGCAGAC 891
Qy 3820 AGCAGGCCCCGAGCCCTTCAGGCCCAAGCAGCTGCCCGCAGACTCGCTGGCACTCAGTTCCC 3879
Db 892 AGCAGGCCCCGAGCCCTTCAGGCCCAAGCAGCTGCCCGCAGACTCGCTGGCACTCAGTTCCC 951
Qy 3880 TCATCTCTAAAGGTG-AAGGGTGTGACGAGGATATGCTTCAGCAGGAACAGTCTGTGGATGG 3938
Db 952 TCATCTCTAAAGGTGAAAGGGTGTGACGAGGATATGCTTCAGCAGGAACAGTCTGTGGATGG 1011
Qy 3939 ACATGATCAGTCTAAAGAAAGCAGCAGAGACGCTCGCGCGCCCGCAGCCCACTATC 3998
Db 1012 ACATGATCAGTCTAAAGAAAGCAGCAGAGACGCTCGCGC-CCMGCCCCCACTATC 1070
Qy 3999 AGTCTCCAGGCTGCTGTTCCCGCAGACAC 4028
Db 1071 AGTCT-CAGAGTCTGTTCCAGGAAGACAS 1099

RESULT 9
LOCUS BM458601
DEFINITION BM458601 1140 bp mRNA linear EST 05-FEB-2002
AGENCOURT_6413524 NIH_MGC_85 Homo sapiens cdna clone IMAGE:5497658
5', mRNA sequence.
ACCESSION BM458601
VERSION BM458601.1 GI:18507641
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1140)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Lou Straudt
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12128 row: c column: 03
High quality sequence stop: 724.

FEATURES

source
1. .1140
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5497658"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 19.1%; Score 876.2; DB 4; Length 1140;
Best Local Similarity 94.1%; Pred. No. 2.1e-209; Mismatches 50; Indels 15; Gaps 11;
Matches 1030; Conservative 0;

QY 2556 AGAATTATTCCTCAACAGCCTGTCCAGCATTCTCTGGAGGAGCAGGACAAATCT 2615
DB 44 AGAATTATTCCTCAACAGCCTGTCCAGCATTCTCTGGAGGAGCAGGACAAATCT 103

QY 2616 CTCGTGCTCAGCATCGACAGCCTCTCCCTGTCGATGACATGAGAGAACCCATCAAA 2675
DB 104 CTCGTGCTCAGCATCGACAGCCTCTCCCTGTCGATGACATGAGAGAACCCATCAAA 163

QY 2676 GGCCTCTCAAGCTCGGGGACACCTGAGCTCAGCGTACACTCTCTGGAGCAGCGC 2735
DB 164 GGCCTCTCAAGCTCGGGGACACCTGAGCTCAGCGTACACTCTCTGGAGCAGCGC 223

QY 2736 CGAGGCTCGAAGCTCAGCTGGAACATGTGTGCGCGCGCGCGCCACCGACACCCC 2795
DB 224 CGAGGCTCGAAGCTCAGCTGGAACATGTGTGCGCGCGCGCGCCACCGACACCCC 283

QY 2796 AAGCTATTTCAATGTGTGAAGTCAAAATACAGTCTCTTAATGTGTAACCTGACAT 2855
DB 284 AAGCTATTTCAATGTGTGAAGTCAAAATACAGTCTCTTAATGTGTAACCTGACAT 343

QY 2856 CCGGAGGTTCCACCGGTTCAAGTGGAGACATCGCCATCGCATCAGCAGCCAGATCCC 2915
DB 344 CCGGAGGTTCCACCGGTTCAAGTGGAGACATCGCCATCGCATCAGCAGCCAGATCCC 403

QY 2916 AGCTCAGCCTTCAGCTTGGTCAACAAAGACGGGAGCCTGTTCTGCTACGATGGAGT 2975
DB 404 AGCTCAGCCTTCAGCTTGGTCAACAAAGACGGGAGCCTGTTCTGCTACGATGGAGT 463

QY 2976 GCAGACTCGGCGATCGACTCGAGTGCACATGCGCCCTGTATGGAGCTTCGCTGGAG 3035
DB 464 GCAGACTCGGCGATCGACTCGAGTGCACATGCGCCCTGTATGGAGCTTCGCTGGAG 523

QY 3036 CTGGAGGTTCAAGCATGGCCAGCTGGAGAACAGGCGCTTAACCTGCCCTCCACCGCGGC 3095
DB 524 CTGGAGGTTCAAGCATGGCCAGCTGGAGAACAGGCGCTTAACCTGCCCTCCACCGCGGC 583

QY 3096 TCCACACTGCCGGAAGACGAGCCTTCTGCTCGGTGCAAGATGCTGCTGCAAAACACAGG 3155
DB 584 TCCACACTGCCGGAAGACGAGCCTTCTGCTCGGTGCAAGATGCTGCTGCAAAACACAGG 641

QY 3156 CTCAGCGGTTCCAGGGATTCGACGCCCTCCAGCTTCAGTGGGAAACAGGCGCTTCGCA 3215
DB 642 CTCAGCGGTTCCAGGGATTCGACGCCCTCCAGCTTCAGTGGGAAACAGGCGCTTCGCA 700

QY 3216 GCAGCAAGTGGGGGACAGCAAGTCCCTCCAGATTTTCACTGAGCTGAGCCCTGCCAC 3275
DB 701 GCAGCAAGTGGGGGACAGCAAGTCCCTCCAGATTTTCACTGAGCTGAGCCCTGCCAC 760

QY 3276 CCTGCTGAAAAACA-TCCGCCACGTGAAGACAGAGGAGGATGGCAGGATTTACCTG 3334

DB 761 CCTGCTGAGAAAACACTCGGCACGTAAGAGACAG-AGGAGGATGGCAGGATTTACCTC 819
QY 3335 GGGAAACAAAACAGGGATCTTTTCTGCTGCTCCAGTCCAGTGGGCTGACCGGCTT 3394
DB 820 GGGGAAAACAAACAGG---ATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 876
QY 3395 GGATCAGTGACCATTTGTTGGCAGACAGGGGAGACAGCTTCCAGCCT-GGGTCAGAAAG 3453
DB 877 GGATCAGTGAGCATTTGCTGGCAGACA-GGAGAGCAGCTTCCAGCTGGGGTCAAAGG 935
QY 3454 GGTGGGCGAG-CCCTTCGCGCCCTCACCTCCAGCTCTGTGAGAGTGTCAAGTGTGTA 3512
DB 936 GGTGGGCGAGCCCTTGGGCGCCCTCACCTCCAGCTGCTGTGAAAGTNTCAAGTGTGTA 995
QY 3513 AGGGCCCAAACTCAGGTTCACTGAGACAGGAGGAGGATGATGCC--CGGCCCTAGGT 3570
DB 996 AGGGCCCAAACTCAGGTTCACTGAGACAGGAGGAGGATGATGCC--CGGCCCTAGGT 1055
QY 3571 TAAAGGGGCGCTCTAAACCCCTTGGCTGCTGCTCACCTGCGCCAGCT--CACCCCTTTTGGG 3628
DB 1056 TAGGGGGGCGCTCTAAACCCCTTGGCTGCTGCTTAACGGNCAGGTTACCCCTTTTGGG 1115
QY 3629 TGTAGGGGAAAGAA 3643
DB 1116 GGAAGGGGAAAGAA 1130

RESULT 10
BU541005
LOCUS
DEFINITION BU541005 955 bp mRNA linear EST 13-SEP-2002
AGENCY: 10326204 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:6572558
5' mRNA sequence.
ACCESSION BU541005
VERSION BU541005.1 GI:22851446
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 955)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaapb-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12128 row: n column: 14
High quality sequence start: 31
High quality sequence stop: 630.
Location/Qualifiers
1. .955
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6572558"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 42"
/note="Organ: pancreas; Vector: pMTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).

FEATURES
source

Note: this is a NIH_MGC Library. | "

ORIGIN
Query Match 17.7%; Score 811.4; DB 5; Length 955;
Best Local Similarity 95.3%; Pred. No. 4.6e-193;
Matches 880; Conservative 0; Mismatches 36; Indels 7; Gaps 4;

Qy 3681 ACACCTTTTCAGGTTGTTCAACACACAGGTCCTGAGTTGACCTCTGTTGTCAGCAAGGACCA 3740
Db 32 ACACCTTTTCAGGTTGTTCAACACACAGGTCCTGAGTTGACCTCTGTTGTCAGCAAGGACCA 91

Qy 3741 AAGAGGTTGTTAAGTGAAGTGGTCTCAGTCCCCAGACATGTCCTTGTGCTGGC 3800
Db 92 AAGAGGTTGTTAAGTGAAGTGGTCTCAGTCCCCAGACATGTCCTTGTGCTGGC 151

Qy 3801 TACCACCTCTTCCACAGACACAGCCCGGAGCCCTTTCAGGCCAGCAGTCCGCCAGAC 3860
Db 152 TACCACCTCTTCCACAGACACAGCCCGGAGCCCTTTCAGGCCAGCAGTCCGCCAGAC 211

Qy 3861 TCGCTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGTGATGCAGGATATGCTGACA 3920
Db 212 TCGCTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGTGATGCAGGATATGCTGACA 271

Qy 3921 GGAACAGTCTGTGATGACATGATCATGCTAAGGTAAGGTAAGGTAAGGTAAGGTAAGG 3979
Db 272 GGAACAGTCTGTGATGACATGATCATGCTAAGGTAAGGTAAGGTAAGGTAAGGTAAGG 331

Qy 3980 GCGCCCGAGCCCACTATCAGTGTCCAGCGTGGTGGTCCCGAGCAGCAGCTCAGCATC 4039
Db 332 GCGCCCGAGCCCACTATCAGTGTCCAGCGTGGTGGTCCCGAGCAGCAGCTCAGCATC 391

Qy 4040 ACACGTGACATCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4099
Db 392 ACACGTGACATCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451

Qy 4100 TCTGATGACCTCAAGACATTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4159
Db 452 TCTGATGACCTCAAGACATTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 511

Qy 4160 ACTGTAGGACATACGACGACGAGATGGGGTGAAGGACACAGTCTTGAGCTGCTCAC 4219
Db 512 ACTGTAGGACATACGACGACGAGATGGGGTGAAGGACACAGTCTTGAGCTGCTCAC 571

Qy 4220 ATGATGTGACTCTCAAACTCTTCCAGATTTCTTAAGATAGCACCCCTTCCCAT 4279
Db 572 ATGATGTGACTCTCAAACTCTTCCAGATTTCTTAAGATAGCACCCCTTCCCAT 631

Qy 4280 TGCCCCAGCTTTCCTTCTCCAGGGGAGTACTCAGGACTCAGTAGCATTAATCA 4339
Db 632 TGCCCCAGCTTTCCTTCTCCAGGGGAGTACTCAGGACTCAGTAGCATTAATCA 691

Qy 4340 CTTGTGATCTCA - GGGGGTGTCTGTAGCTCACTCTGCGGAGGGGAGCGCGAG 4398
Db 692 CTTGTGATCTCA - GGGGGTGTCTGTAGCTCACTCTGCGGAGGGGAGCGCGAG 751

Qy 4399 ACTCGTGGGAGAGCTCATTTCCACATCTTTGCCAAGACAGCTTTTGTCCAGCTGCTCAC 4458
Db 752 ACTCGTGGGAGAGCTCATTTCCACATCTTTGCCAAGACAGCTTTTGTCCAGCTGCTCCA 811

Qy 4459 ATTGAGTCAGACTGCTCCCGGGAGAGAGCCCGGCCCCAGCAGCATAAAGAACTGCGAGC 4518
Db 812 CATGAGTCAGACTGCTCCCGGGAGAGAGCCCGGCCCCAGCAGCATAAAGAACTGCGAGC 871

Qy 4519 CTTGG - - - - TACTGAGAGTCTGGGTGTAGAGAACTC - TTTGTAAAGCAATAAGTTGG 4573
Db 872 CTTGGTACTGCAAAAGTCTGGGGTGTGAAGAGAACTCTTTTGTAGCCACTAAAGTTGGG 931

Qy 4574 GGTGATGACAAATGTTAAAAA 4596
Db 932 GGTGATGACAAATGTTAAAAA 954

RESULT 11

BM454415
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BM454415 1038 bp mRNA linear EST 05-FEB-2002
AGENCOURT 6402274 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5497584
5', mRNA sequence.
BM454415
BM454415.1 GI:18503455
EST.
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 1038)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rcs@nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12127 row: p column: 01
High quality sequence stop: 595.

FEATURES
source

1..1038
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5497584"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/note="Origin: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 17.6%; Score 808.8; DB 4; Length 1038;
Best Local Similarity 96.2%; Pred. No. 2.1e-192;
Matches 883; Conservative 0; Mismatches 27; Indels 8; Gaps 5;

Qy 630 TTGGAAGGAAAGCGTCCGAGCAAGCCCGGAGAAACGGAAGAGAGAGCTCAAGTC 689
Db 2 TTGGAAGGAAAGCGTCCGAGCAAGCCCGGAGAAACGGAAGAGAGAGCTCAAGTC 61

Qy 690 CCTGGCTCATGAGGAGTGGCTTGGCCAAACCCCTCCAGGACCCCTGAGCAGGAGAG 749
Db 62 CCTGGCTCATGAGGAGTGGCTTGGCCAAACCCCTCCAGGACCCCTGAGCAGGAGAG 121

Qy 750 CTGCACCATCCAGTCAGGAGGATGAGTCTCCACTCGGCGCCCATATGTTAGAAACAC 809
Db 122 CTGCACCATCCAGTCAGGAGGATGAGTCTCCACTCGGCGCCCATATGTTAGAAACAC 181

Qy 810 CCAGCAGTTCCCAAGCCTCTGAAGGAAACAGGCTTGGGCAACTCTGTTTAAAGCAGCT 869
Db 182 CCAGCAGTTCCCAAGCCTCTGAAGGAAACAGGCTTGGGCAACTCTGTTTAAAGCAGCT 241

Qy 870 TGGCAGGGCTACGGCGGCTCTGCTCGATCAGAACTCCACAACTGATCAGCCCTT 929
Db 242 TGGCAGGGCTACGGCGGCTCTGCTCGATCAGAACTCCACAACTGATCAGCCCTT 301

Qy 930 GCATGTCTGAACACACCTGTGGAACTGCACACCCCGAGGAGGAGCCCTTGGCCCT 989
Db 302 GCATGTCTGAACACACCTGTGGAACTGCACACCCCGAGGAGGAGCCCTTGGCCCT 361

Qy 990 GCCCAGCAGCCCTTCCCTTATAGCAGACTGCTCCTTCCCTTCCCTTCCCTTCCCTTCCCA 1049
Db 362 GCCCAGCAGCCCTTCCCTTATAGCAGACTGCTCCTTCCCTTCCCTTCCCTTCCCTTCCCA 421

1050 GGCCTGAAACCTCACCTCTTGAGTCCTTCTTGGGCAAACTGGCCTGTGTAGACAGCCA 1109
Db GCCTGGAACCTCACCTCTTGAGTCCTTCTTGGGCAAACTGGCCTGTGTAGACAGCCA 481
1110 GAAACCTTTGCCCTGACCCACACCTGAGCAAACTGGCCTGTGTAGACAGTCCTCAAGCCCT 1169
Db GAAACCTTTGCCCTGACCCACACCTGAGCAAACTGGCCTGTGTAGACAGTCCTCAAGCCCT 541
1170 GCCTGGCCACACCTTGAGCCACAGTCGCTGTCTGTGTGGCCCATGAGAACTTTTCTGT 1229
Db GCCTGGCCACACCTTGAGCCACAGTCGCTGTCTGTGTGGCCCATGAGAACTTTTCTGT 601
1230 CGAGGAATACCTAGTCATGCTCTGCAAGCGAGCTGAGCTCAAGCCAGGCCACAGCCT 1289
Db CGAGGAATACCTAGTCATGCTCTGCAAGCGAGCTGAGCTCAAGCCAGGCCACAGCCT 661
1290 GACCACTGGCCGCAAGACTGGGCGACACCGGGCTCCAGATCCCGGAGCCAGCCCA 1349
Db GACCACTGGCCGCAAGACTGGGCGACACCGGGCTCCAGATCCCGGAGCCAGCCCA 721
1350 AACTGAGGACACGAGG-TGTCCTGCTCACTGAGAACTCAAGCCAG-TGATATGAG 1407
Db AACTGAGGACACGAGG-TGTCCTGCTCACTGAGAACTCAAGCCAG-TGATATGAG 781
1408 TACCGAGAAAGTCCACTGGGCGACGACCACTCGGCTCGGCGAGAGGCT-CCTTCGG 1466
Db TACCGAGAAAGTCCACTGGGCGACGACCACTCGGCTCGGCGAGAGGCT-CCTTCGG 841
1467 AGAGTGTGAC--AGGATGAGGACCAAGCAGACTGGC--TTCCAGTGGCTGTCAAAAAG 1521
Db AGAGGCGCACCAAGATGGGAGACCAAGCAGACTGGCCTTCAGTGGCCTGTCCAAAAG 901
1522 GTGCGCTGAGATTT 1539
Db GTGCGCTGAGAAAT 919

RESULT 12
BX331085/c
LOCUS
DEFINITION BX331085 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DK011YM10 3-PRIME, mRNA sequence.
ACCESSION BX331085
VERSION BX331085.2 GI:46263202
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 909)
Li W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 1, 2003 this sequence version replaced gi:30307969.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
1. .909
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011YM10"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 17.1%; Score 787.4; DB 5; Length 909;
Best Local Similarity 95.7%; Pred. No. 5.1e-187; Mismatches 33; Indels 6; Gaps 6;
Matches 872; Conservative 0;
QY 3679 CCACACTTTTTCAGGTGTTGCAACACA-GGTCCTGAGTTG-ACCTCTGTTTCAGCAAGG 3736
Db CCACGCTTTTCAGGTTAGTACAAACAGGGTCTGAGTTGAACCTCTGTTCCAGCAAGG 850
QY 3737 ACCAAGAAGGTGTAGTGAAGTGTTCAGTCCCT-AGACATGTGCCCTTTGCTG 3795
Db GCCCAAAAAAGGTGTGAATGAAATGGGTTTTCAGTCCCAAGACATGGGCCCCCTTGGCTG 790
QY 3796 CTGGCTTACACTCTTTCCTCCAGAGCAGAGGCCCGAGGCCCTTTCAGGGCCAGCACTGGCC 3855
Db CGGGCTACCACTCTTTCCTCCAGAGCAGAGGCCCGAG-CCCTTCAGGCCAGGCATGGCT 731
QY 3856 CAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGTGAAGGTGATGCAAGGATATGCC 3915
Db CAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGTGAAGGTGATGCAAGGATATGCC 671
QY 3916 TGACAGGAACAGTCTGTGGATGGAATCATCATGTCTTAAGGAAAGCAGCAGAGAGAG 3975
Db TGACAGGAACAGTCTGTGGATGGAATCATCATGTCTTAAGGAAAGCAGCAGAGAGAGAG 611
QY 3976 -TCCGGGCCCCCAGGCCCACTATCAGTGTCCAGCGTGTCTTCCAGAGCAGAGCTCA 4034
Db CTCGGGCCCCCAGGNCACATCATCATGTCTCCAGCGTGTCTTCCAGAGCAGAGCTCA 551
QY 4035 GCATCACACTGACACTCACTCCCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 4094
Db GCATCACACTGACACTCACTCCCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 491
QY 4095 TGCACCTGTGATGACCTCAAGCAGCTTTTCATGGCTGCGCTGCGCTGCGCTGCGCTGCGCT 4154
Db TGCACTCTGTGATGACCTCAAGCAGCTTTTCATGGCTGCGCTGCGCTGCGCTGCGCTGCGCT 431
QY 4155 GTGACACTGTAGGAGCATAGCAAGCCAGAGATGGGTGAAGGGACACAGTCTTTGAGCTG 4214
Db GTGACACTGTAGGAGCATAGCAAGCCAGAGATGGGTGAAGGGACACAGTCTTTGAGCTG 371
QY 4215 TCCACATGATGATGATCTCTCAAACTCTTCCAGATTTCTTAAGATAGCAACCCCTTC 4274
Db TCCACATGATGATGATCTCTCAAACTCTTCCAGATTTCTTAAGATAGCAACCCCTTC 311
QY 4275 CCCATTGCCAGCTTAGCCTCTTCCAGGGGAGTCTTCCAGGAGTCTCAGGACTCAGTAGCATTA 4334
Db CCCATTGCCAGCTTAGCCTCTTCCAGGGGAGTCTTCCAGGAGTCTCAGGACTCAGTAGCATTA 251
QY 4335 AATCAGCTGTGAATCGTTCAGGGGTGTCTGTAGCTCAACCTCTCTGGGGCAGGGGAGCG 4394
Db AATCAGCTGTGAATCGTTCAGGGGTGTCTGTAGCTCAACCTCTCTGGGGCAGGGGAGCG 191
QY 4395 CGAGACTCCGTGGGAGAGCTCATTTCCACATCTTTCGCAAGACAGCCTTTTGTCCAGCTGT 4454
Db CGAGACTCCGTGGGAGAGCTCATTTCCACATCTTTCGCAAGACAGCCTTTTGTCCAGCTGT 131
QY 4455 CCACATTGAGTCAGACTGCTCCCGGGAGAGAGCCCGGCCCGCCAGCATATAAGAACTG 4514
Db CCACATTGAGTCAGACTGCTCCCGGGAGAGAGCCCGGCCCGCCAGCATATAAGAACTG 71
QY 4515 CAGCCTTGGTACTGCAGAGTCTGGGTGTAGAGAACTTTTGTAAAGCAATAAAGTTTGGG 4574
Db CAGCCTTGGTACTGCAGAGTCTGGGTGTAGAGAACTC-TTGTAAAGCAATAAAGTTTGGG 12
QY 4575 GTGATGACAA 4585
Db GTGATGACAA 1


```
RESULT 13
BG775354/c
LOCUS      602649808T1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4761048 3', linear mRNA 1037 bp EST 15-MAY-2001
DEFINITION mRNA sequence.
ACCESSION  BG775354
VERSION     BG775354.1 GI:14045671
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1037)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: DCTD/BTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LICM1612 row: o column: 01
            High quality sequence start: 2
            High quality sequence stop: 864.
            Location/Qualifiers
                1..1037
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4761048"
                /tissue_type="carcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH MGC 40"
                /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
                Site 2: EcoRI; cDNA made by oligo-dT priming.
                Directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCACGAG(G). Library constructed by
                Ling Hong in the laboratory of Gerald M. Rubin (University
                of California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: this is a NIH_MGC Library."
FEATURES
    source
ORIGIN
Query Match      16.5%; Score 756.2; DB 4; Length 1037;
Best Local Similarity 93.3%; Pred. No. 4e-179;
Matches 889; Conservative 0; Mismatches 53; Indels 11; Gaps 9;

Qy 3588 CCCCTTGCTGGCTCCTGCTGCTCACTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAAGATGCC 3647
Db 949 CTTCTAACCCCTGCTGGCCACACAGCGGAGCTACCCCTTTTGGTCAGGGAACATGAACGCC 890

Qy 3648 TGACCTGGGAGGCTCCCTGGTAGATACACCACTTTTCAGGTTGTTGCACACAGG 3707
Db 889 TGAACCCGGGAG--CTCCCTGTAGATACACCACTTTTCA--GTTGTTGCAACACAGG 833

Qy 3708 TCCTGAGTTGACCTCTGTTTCA--GCCAAGGACCAAGAGGTGTAGTGAAGTGGTTTC 3766
Db 832 TCTGAGTTGGACTCCCGTTTCAGCCAGGACCAAGAGGTGTGTA--TGAATGGTTTC 775

Qy 3767 TCAGTCCCAACATATG--CCCTTTTGTGCT--GGCTACCACTCTTCCCCAGACAGCAG 3824
Db 774 TCAGTCCCAACAGAGCGCCCTTGTGCTGGGCTACCACTCTTCCCCAGACAGCAG 715

Qy 3825 GCCCGAGCCCTTCAGGCCAGCACTGCCAGACTCGCTGGCACTCAGTTCCCTCATC 3884
Db 714 GCCCGAGCCCTTCAGGCCAGCACTGCCAGACTCGCTGGCACTCAGTTCCCTCATC 655
```

```
Qy 3885 TGTAAAGGTGAAGGGTGATGTCAGGATATGCTCTGA CAGGAACAGTCTGTGGATGACATGA 3944
Db 654 TGTAAAGGTGAAGGGTGATGTCAGGATATGCTCTGA CAGGAACAGTCTGTGGATGACATGA 595

Qy 3945 TCAGTGTCTAAGGAAGCAGCAG-AGAGAGAG--TCCGCGCCGCCAGCCCCCACTATCAGTG 4002
Db 594 TCAGTGTCTAAGGAAGCAGCAGTAGAGAGCGTCCGCGCCGCCAGCCCCCACTATCAGTG 535

Qy 4003 TCCAGCGTGTGTTTCCCCAGAGCAGACAGCTCAGCATCACACTGACACTCACCCCTGCCCTG 4062
Db 534 TCCAGCGTGTGTTTCCCCAGAGCAGACAGCTCAGCATCACACTGACACTCACCCCTGCCCTG 475

Qy 4063 CCCCTGGCCAGAGGGTACTGCCGACGGACACTTTTGCACCTGTGATGACCTCAAAGCACTTTC 4122
Db 474 CCCCTGGCCAGAGGGTACTGCCGACGGACACTTTTGCACCTGTGATGACCTCAAAGCACTTTC 415

Qy 4123 ATGCTCCCTCTCTGCGAGGCGAGGCGAGGAGACACTGTAGGAGCATAGCAAGCCAG 4182
Db 414 ATGCTCCCTCTCTGCGAGGCGAGGCGAGGAGGAGGAGTACACTGTAGGAGCATAGCAAGCCAG 355

Qy 4183 GAGATGGGGTGAAGGAGCACACAGTCTTGAGCTGTCCACATGATGTGACTCTCTCAAACTC 4242
Db 354 GAGATGGGGTGAAGGAGCACACAGTCTTGAGCTGTCCACATGATGTGACTCTCTCAAACTC 295

Qy 4243 TTCAGATTTCTTAAAGATAGACACCCCTTCCCATTTGCCCACTGCTTTCTTCC 4302
Db 294 TTCAGATTTCTTAAAGATAGACACCCCTTCCCATTTGCCCACTGCTTTCTTCC 235

Qy 4303 CAGGGAGCTACTCAGGACTCAGTAGCATTAATAGCTGTGAATCGTCAGGGGGTGTTC 4362
Db 234 CAGGGAGCTACTCAGGACTCAGTAGCATTAATAGCTGTGAATCGTCAGGGGGTGTTC 175

Qy 4363 TGCTAGCTCAACCTCTCTGGGGCAGGGGAGCCCGAGACTCCGTTGGGAGAACTCATTTCCC 4422
Db 174 TGCTAGCTCAACCTCTCTGGGGCAGGGGAGCCCGAGACTCCGTTGGGAGAACTCATTTCCC 115

Qy 4423 ACATTTTCCCAAGACACCTTTTGTCCAGCTGTCCACATTTGATGAGTCAGAGTGTTC-CCGGGG 4481
Db 114 ACATTTTCCCAAGACACCTTTTGTCCAGCTGTCCACATTTGATGAGTCAGAGTGTTCACCGGGG 55

Qy 4482 AGAGAGCCCCGGCCCCCAGCACATAAAGAACTGGAGCTTGGTACTCGAGGT 4534
Db 54 AGAGAGCCCCGGCCCCCAGCACATAAAGAACTGGAGACTTGGTACTCGAGGT 2

RESULT 14
BG822015
LOCUS      602726122F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4865789 5', linear mRNA sequence.
DEFINITION mRNA sequence.
ACCESSION  BG822015
VERSION     BG822015.1 GI:14169602
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 774)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: NIH Intramural Sequencing Center
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LICM1731 row: k column: 06
            High quality sequence stop: 748.
            Location/Qualifiers
FEATURES
```


source	1. .774		/organism="Homo sapiens"		RESULT 15	
	/mol_type="mRNA"		/db_xref="taxon:9606"		AK038828	
	/clone="IMAGE:4865789"		/issue_type="adenocarcinoma cell line"		LOCUS	
	/lab_host="DH10B (phage-resistant)"		/clone_lib="NIH_MGC_15"		DEFINITION	
	/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"				ACCESSION	
					VERSION	
					KEYWORDS	
					SOURCE	
					ORGANISM	
					REFERENCE	
					AUTHORS	
					TITLE	
ORIGIN					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	
					REFERENCE	
					AUTHORS	
					TITLE	
					JOURNAL	
					MEDLINE	
					PUBMED	

Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>
Location/Qualifiers
source
1. .2964
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:A230066D12"
/db_xref="taxon:10090"
/clone="A230066D12"
/sex="male"
/tissue_type="hypothalamus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
1. .2964
/note="mitogen-activated protein kinase kinase 14
(MGP|MGI:1858204, GB|NM_016896, evidence: BLASTN, 100%,
match=2829)"
ORIGIN
Query Match 16.1%; Score 741.6; DB 3; Length 2964;
Best Local Similarity 82.6%; Pred. No. 2.7e-175;
Matches 886; Conservative 0; Mismatches 179; Indels 7; Gaps 3;
Qy 329 GGAAGAAACAGAGCTCGCTTACAGCTTGAGCGGTGAGGAGAGAGCGCTGTGTCGCG 388
Db 1 GTAAGAGCAGAGCTGCATCTTCAAGCTTGAGCGGTGAGGAGAGAGCGCTGTGTCG 60
Qy 389 GAAAGTGGAGATCTGATCAGCTGATTACCAAGGCGCAGCCCAAGAGGCTCCGAGG 448
Db 61 GGAAGTGGAGATCTTAAACAGCTGATACCAAGGCGCAGCCCAAGAGGCTCCGAGG 120
Qy 449 CAGGCGCAGCTGCGCATCTTATCGCCAGGCTGAGTGAGAGTACAGAGGTTCA 508
Db 121 GAGAGACACCGCGCATCTTCATCATCGCCAGGCTGATGTGAGAGTACAGAGTTCA 180
Qy 509 GCGCCACCTTTTCAGAACGCAATTTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGA 568
Db 181 GCCCACCTCTCAGAGCGCATTTTCATCGCGGGTCAACAGCTACAGCCAGTCTGAGA 240
Qy 569 GTCTTGATCAGATCCCAAGTGTGCGCCATGTCTACAGAGGCAAAATGSCCGCTGTGT 628
Db 241 GTCTCGATCAAAATCCCAAGTGTGCGCCATGTCAACTGGAAGGCAAAATGSCCGCTGTGT 300
Qy 629 GTTGGAGGGAAGCGTCGACAGCAAGCGCGGAAGAACGGAAGAGAGCTCAAAAT 688
Db 301 GCCGGAGGGAAGACGTACGCGAAGAGCCCGGAAGAACGTAAGAGAGGTCGAAGT 360
Qy 689 CCCTGGGTCAATCAGGAGTGGCTTTGGCCAAACCCCTCCCGAGGACCCCTCAGCAGGAGA 748
Db 361 CACTGGCCCGCAGGAGTGGCTTTAGCCAAAGCCCTGCCCAGAAACCCCTCAGCAGGAGA 420
Qy 749 GCTGACCATCCAGTGCAGGAGGATGATCTCACTCGCGGCGCCCATATGTTTAAAGAAC 808
Db 421 GCTGTACCATCCAGTACAGGAAGATGATCTCACTAGGCAACCTCTATGCCAGAAATG 480
Qy 809 CCCGCGAGTTCACCAAGCCTCTGAAGAACAGGCGCTTGGCAACTCTGTTTAAAGGAGC 868
Db 481 TCTCCAGTTTACCAAGCCTCTGGGGGACACAGGCTTGGCCACCTGTGCTTTAAGAAAC 540
Qy 869 TTGCGAGGCGCTACGGCGGCTCTGCGCTCGATCAGAACTCCAAACTGATCAGCCCT 928
Db 541 AAGATGAAGGCTCGGACCGTACTGCTCGACAGAACTCCAAACTGATCAGCCCT 600
Qy 929 TGCATGTGTGAACCAAGTGTGGAACACTGCACACCCCGCAGGACGGAGGCGCCCTGCCCC 988
Db 1021 GGCCCCGAAACTGAGGAGAACGAGGGGGTCTCTGCTTACTGAGGTAACTAACC 1072

601 TGCATGTCTTAAACACAGCTGTGGAAACTCCACACACCCAGGCCACAGGCCCGGCCCC 660
989 TGCCC-ACGACACCCCTTCCCTATAGCAGACTGCTCATCCCTTCCATTCCACCCCTCTC 1047
661 CACCGGACTCACCCCTTCCCTTACAGGGAATGCCCATCCTTTCCCATTTACCCCTG 720
1048 CAGCCCTGGAAACCTCACCCCTCTGGAGT---CCTTCTGGGCAACTGGCTGTGTAGAC 1104
721 GAGCCCTGGAAACCCCTATATGCTGGACTCTGCCGTCTTGACAAACTAGCCGGTGTGAGC 780
1105 AGCCAGAAACCCCTTGCCTGACCCCA---CACCTGAGCAAACTGGGCTGTGTAGACAGTCCA 1161
781 GGCCAGCGGCTCTGCTGCTGCGCCACCCGATCTAAGCCAACTGGCCCATGGAGACAGTCAG 840
1162 AAGCCCTGCTGCTGCGCCACACCTGGAGCCAGCTGCTGTCTGTGTGTCGCTATGAGAG 1221
841 AAGCCGCTGCTGCGCCACACCTGGAGTCCAGCTGCCGCTCTCGGGGTGCCCTAGAAAAAG 900
1222 TTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGCAGCGTGAAGTCAAGCCAGGCC 1281
901 GTTCCCGTGGAGGAATACCTGGTGCATGCGCTCCAAGGAAGTGTGAGCTCAGGCCAGGCC 960
1282 CACAGCCTGACCCAGCCTGGCCAAAGACCTGGGCGACACGGGCTCCAGATCCCGGAGGCC 1341
961 CACAGCCTGGCCAGCCTGGCTAAGACATGTCTCTCGGAAGCGCCAAAGCTCAGAGGCTC 1020
1342 AGCCCCAAAACCTGAGGACAAACGAGGGTGTCTCTGCTCACTGAGAAACTCAAGC 1393
1021 GGCCCCGAAACTGAGGAGAACGAGGGGGTCTCTGCTTACTGAGGTAACTAACC 1072

Search completed: June 10, 2005, 20:27:55
Job time : 9696.08 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 14:23:14 ; Search time 1746.62 Seconds

(without alignments)
16311.995 Million cell updates/sec

Title: US-09-155-676B-6

Perfect score: 4596

Sequence: 1 agcggggggactgtgccgtg.....gatgacaaatgttaaaaaa 4596

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 309930249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4578.4	99.6	4596	10	US-09-981-397A-17
2	4578.4	99.6	4596	17	US-10-172-118-975
3	4578.4	99.6	4596	18	US-10-342-887-975
4	4578.4	99.6	4596	19	US-10-283-975A-294
5	4440	96.6	4684	13	US-10-087-192-887
6	2837.6	61.7	2844	9	US-09-871-889-2
7	2042.6	44.4	2828	13	US-10-087-192-884

8	1556.2	33.9	73967	13	US-10-087-192-886	Sequence 886, App
9	1304	28.4	1385	18	US-10-296-115-97	Sequence 97, Appl
10	480.2	10.4	42566	13	US-10-087-192-883	Sequence 883, App
11	349.4	7.6	2475	17	US-10-108-260A-1754	Sequence 1754, App
12	291.6	6.3	6673	19	US-10-433-793-141	Sequence 141, App
13	267.4	5.8	499	10	US-09-918-995-2585	Sequence 2585, Ap
14	243.6	5.3	6673	19	US-10-433-793-142	Sequence 142, App
15	220.6	4.8	296	9	US-09-783-590-11852	Sequence 11852, A
16	171.8	3.7	500	16	US-10-029-386-8637	Sequence 8637, Ap
17	166.8	3.6	293	16	US-10-029-386-22365	Sequence 22365, A
18	88.8	1.9	2348	18	US-10-380-127A-3	Sequence 3, Appli
19	88.4	1.9	455	20	US-10-357-930-22052	Sequence 22052, A
20	88.4	1.9	455	20	US-10-357-930-27909	Sequence 27909, A
21	88.4	1.9	476	20	US-10-357-930-12325	Sequence 12325, A
22	88.4	1.9	504	20	US-10-357-930-33496	Sequence 33496, A
23	88.4	1.9	504	20	US-10-357-930-42420	Sequence 42420, A
24	83.8	1.8	3332	13	US-10-000-864-11	Sequence 11, Appl
25	82.2	1.8	1935	13	US-10-000-864-5	Sequence 5, Appli
26	73.8	1.6	463	20	US-10-357-930-3156	Sequence 3156, Ap
27	66.2	1.4	1300	18	US-10-425-114-30892	Sequence 30892, A
28	66.2	1.4	1485	20	US-10-425-115-111586	Sequence 111586,
29	63.8	1.4	825	17	US-10-369-493-26829	Sequence 26829, A
30	62.6	1.4	1254	17	US-10-438-339-2	Sequence 2, Appli
31	62.6	1.4	1254	18	US-10-416-477-2	Sequence 2, Appli
32	62.6	1.4	1254	21	US-10-754-829A-2	Sequence 2, Appli
33	62.6	1.4	1278	18	US-10-425-114-26294	Sequence 26294, A
34	62.6	1.4	1286	18	US-10-425-114-26222	Sequence 26222, A
35	62.6	1.4	1318	18	US-10-425-114-26279	Sequence 26279, A
36	62.6	1.4	2419	17	US-10-108-260A-974	Sequence 974, App
37	62.6	1.4	2626	18	US-10-433-794-32	Sequence 32, Appl
38	62	1.3	4494	18	US-10-263-929-4	Sequence 4, Appli
39	61.6	1.3	3260	9	US-09-858-754-1	Sequence 1, Appli
40	61.6	1.3	3260	19	US-10-728-019-1	Sequence 1, Appli
41	61.6	1.3	5253	9	US-09-858-754-13	Sequence 13, Appl
42	61.6	1.3	5253	13	US-10-000-864-7	Sequence 7, Appli
43	61.6	1.3	5253	19	US-10-728-019-13	Sequence 13, Appl
44	61.4	1.3	1984	18	US-10-425-114-19039	Sequence 19039, A
45	61.4	1.3	2062	18	US-10-425-114-34609	Sequence 34609, A

ALIGNMENTS

RESULT 1

US-09-981-397A-17
; Sequence 17, Application US/09981397A
; Publication No. US20030082519A1
; GENERAL INFORMATION:
; APPLICANT: Axixma Pharmaceuticals AG
; APPLICANT: Schubart, Daniel
; APPLICANT: Habenberger, Peter
; APPLICANT: Stein-Gerlach, Matthias
; APPLICANT: Bevec, Dorian
; TITLE OF INVENTION: Cellular Kinases Involved in Cytomegalovirus Infection and their
; TITLE OF INVENTION: Inhibition
; FILE REFERENCE: AXM-004.1 US
; CURRENT APPLICATION NUMBER: US/09/981.397A
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/240,750
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-981-397A-17

Query Match 99.6%; Score 4578.4; DB 10; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 5 GGGGGACTGTGCCGTGTGGAACGTGTAGCTGTGGA-AGGTGGACTCTGTATTACATTGAGG 63

Db 5 GGGGACATGTGCGCGTGTGAAACGTGTAGTGTGTAGAGGTGACATCTCTTTACCAATTTGAGG 64
Qy 64 ATGTTTGGAGGATGAGTATGTGTGGCAGAGGCACACATAAAACAGGACAGAGACCCCTTTGCC 123
Db 65 ATGTTTGGAGGATGAGTATGTGTGGCAGAGGCACACATAAAACAGGACAGAGACCCCTTTGCC 124
Qy 124 CTTGCTCTTTCCTCCCAACCCAGGCTGACCTGTGTTCTCCCGAGTCTGGGATTTCTAAGT 183
Db 125 CTTGCTCTTTCCTCCCAACCCAGGCTGACCTGTGTTCTCCCGAGTCTGGGATTTCTAAGT 184
Qy 184 GACCTGCTCTGTGTTGTTCTCTCAGATGAGCACAAAGCTGGGAGATGGCAGTGTATG 243
Db 185 GACCTGCTCTGTGTTGTTCTCTCAGATGAGCACAAAGCTGGGAGATGGCAGTGTATG 244
Qy 244 GAAATGGCTGCTCCAGGTGCCCTTGGCTCAGCAGTGGGGCAGCAAGAACTCCCCAAAG 303
Db 245 GAAATGGCTGCTCCAGGTGCCCTTGGCTCAGCAGTGGGGCAGCAAGAACTCCCCAAAG 304
Qy 304 CCAAAAGGAGAGACCGCGCACTGGGGAAGAAACAGAGCTCGGTCTACAAGCTTGAGGCC 363
Db 305 CCAAAAGGAGAGACCGCGCACTGGGGAAGAAACAGAGCTCGGTCTACAAGCTTGAGGCC 364
Qy 364 GTGGAGAGAGCCCTGTGTTCTGCGAAAGTGGGAGATCCTGAATGACGTGATTACCAAG 423
Db 365 GTGGAGAGAGCCCTGTGTTCTGCGAAAGTGGGAGATCCTGAATGACGTGATTACCAAG 424
Qy 424 GGCACAGCCAAAGAGGCTCCAGGCGAGGCCAGCTGCGATCTTATCATCGCCCAAGCT 483
Db 425 GGCACAGCCAAAGAGGCTCCAGGCGAGGCCAGCTGCGATCTTATCATCGCCCAAGCT 484
Qy 484 GAGTGTGAGAAATAGCCAGAGTTACGCCCCACCTTTTCAGAACGCAATTTTCATCGCTGGG 543
Db 485 GAGTGTGAGAAATAGCCAGAGTTACGCCCCACCTTTTCAGAACGCAATTTTCATCGCTGGG 544
Qy 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCCATGCT 603
Db 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCCATGCT 604
Qy 604 ACAGGGGCAAAATGCCCCTGTGTGTGGAAAGGAAAGCGTCGAGCAAAAGCCCGGAAG 663
Db 605 ACAGGGGCAAAATGCCCCTGTGTGTGGAAAGGAAAGCGTCGAGCAAAAGCCCGGAAG 664
Qy 664 AAACGGAAGAGAGAGCTCAAGTCCCTGGCTCATGACAGGAGTGGCTTTGGCCAAACCC 723
Db 665 AAACGGAAGAGAGAGCTCAAGTCCCTGGCTCATGACAGGAGTGGCTTTGGCCAAACCC 724
Qy 724 CTTCCCGAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGCAGGAGATGAGTCTCA 783
Db 725 CTTCCCGAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGCAGGAGATGAGTCTCA 784
Qy 784 CTTGGCGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCTCTGAAAGNACCGAGC 843
Db 785 CTTGGCGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCTCTGAAAGNACCGAGC 844
Qy 844 CTTGGCGCACTCTGTTTAAAGAGCTTGGCGAGGCTTAGCGGCGCTCTGCTCGATCA 903
Db 845 CTTGGCGCACTCTGTTTAAAGAGCTTGGCGAGGCTTAGCGGCGCTCTGCTCGATCA 904
Qy 904 GAACTCCACAAACTGATCAGCCCTTGTGCAATGTCTGAACCAAGTGTGGAACCTGCAACAC 963
Db 905 GAACTCCACAAACTGATCAGCCCTTGTGCAATGTCTGAACCAAGTGTGGAACCTGCAACAC 964
Qy 964 CCCAGGACGAGGCCCCCTGCCCCCTGCCAGCAGCCCTTCCCTTATAGCAGCTGCT 1023
Db 965 CCCAGGACGAGGCCCCCTGCCCCCTGCCAGCAGCCCTTCCCTTATAGCAGCTGCT 1024
Qy 1024 CATCCCTTCCCATTCACACCTCTCCAGCCCTGGAACCTCACCCCTCTGAGTCTCTTCTG 1083
Db 1025 CATCCCTTCCCATTCACACCTCTCCAGCCCTGGAACCTCACCCCTCTGAGTCTCTTCTG 1084
Qy 1084 GGCAAACTGCGCTGTGTAGACAGCCAGAAACCCCTTGCTGACCCACACCTGAGCAACTG 1143

Db 1085 GGCAAACTGCGCTGTGTAGACAGCCAGAAACCCCTTGCTGACCCACACCTGAGCAAACTG 1144
Qy 1144 GCTGTGTAGACAGTCCAAAGCCCCCTGCGCCACACCTGAGGCCACCTGCTGTCT 1203
Db 1145 GCTGTGTAGACAGTCCAAAGCCCCCTGCGCCACACCTGAGGCCACCTGCTGTCT 1204
Qy 1204 CTTGTGTCCCATGAGAAATTTCTGTGGAGGAATACTAGTGCATGCTCTGCAAGGCAGC 1263
Db 1205 CTTGTGTCCCATGAGAAATTTCTGTGGAGGAATACTAGTGCATGCTCTGCAAGGCAGC 1264
Qy 1264 GTGAGCTCAAGCCAGGCCACAGCTGACAGCTGGCAGAGCTGGGAGGAGCCGGGC 1323
Db 1265 GTGAGCTCAAGCCAGGCCACAGCTGACAGCTGGCAGAGCTGGGAGGAGCCGGGC 1324
Qy 1324 TCCAGATCCCGGAGGCCAGCCCAAACTGAGGACAAACAGAGGTGCTGCTCACTGAG 1383
Db 1325 TCCAGATCCCGGAGGCCAGCCCAAACTGAGGACAAACAGAGGTGCTGCTCACTGAG 1384
Qy 1384 AAACCTCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCCACGACAGCTC 1443
Db 1385 AAACCTCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCCACGACAGCTC 1444
Qy 1444 CGCTGGGCAAGGCTCCTTCGGAGAGTGCACAGGATGGAGACAGAGCTGGGCTTC 1503
Db 1445 CGCTGGGCAAGGCTCCTTCGGAGAGTGCACAGGATGGAGACAGAGCTGGGCTTC 1504
Qy 1504 CAGTGGCTGTCAAAAAGGTGCGCTCGAAATATTTCCGGCAGAGAGCTGATGGCATGT 1563
Db 1505 CAGTGGCTGTCAAAAAGGTGCGCTCGAAATATTTCCGGCAGAGAGCTGATGGCATGT 1564
Qy 1564 GCAGGATTGACCTCACCCAGAAATTTGCTTGTATGGAGCTGTGAGAAAGGCGCTTGG 1623
Db 1565 GCAGGATTGACCTCACCCAGAAATTTGCTTGTATGGAGCTGTGAGAAAGGCGCTTGG 1624
Qy 1624 GTCAACATCTTCATGAGAGTCTGGAAGGTGGCTCCCTGGCCAGAGCTGTTCAAGAGCAG 1683
Db 1625 GTCAACATCTTCATGAGAGTCTGGAAGGTGGCTCCCTGGCCAGAGCTGTTCAAGAGCAG 1684
Qy 1684 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGAGCCCTGGAGGCTTGAA 1743
Db 1685 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGAGCCCTGGAGGCTTGAA 1744
Qy 1744 TACCTCCACTCAGAAAGGATTTCTGCATGGGACAGTCAAAAGCTGACACAGTCTCTGTCC 1803
Db 1745 TACCTCCACTCAGAAAGGATTTCTGCATGGGACAGTCAAAAGCTGACACAGTCTCTGTCC 1804
Qy 1804 AGCGATGGGACCGCAGAGCCCTCTGTGACCTTTGGCCATGCTGTGTCTTCAACCTGAT 1863
Db 1805 AGCGATGGGACCGCAGAGCCCTCTGTGACCTTTGGCCATGCTGTGTCTTCAACCTGAT 1864
Qy 1864 GGCTTGGGAAAGTCCCTTGTCTCACAGGGGACTACATCCCTGGCAGACAGACCCATGGCT 1923
Db 1865 GGCTTGGGAAAGTCCCTTGTCTCACAGGGGACTACATCCCTGGCAGACAGACCCATGGCT 1924
Qy 1924 CCGAGGTGTGTCTGGGACAGAGCTGCCACGCGAGGTGGATGTCTGAGCAGCTGCTGT 1983
Db 1925 CCGAGGTGTGTCTGGGACAGAGCTGCCACGCGAGGTGGATGTCTGAGCAGCTGCTGT 1984
Qy 1984 ATGATGCTGCATGCTCAACGGCTGCGCCCTGAGTCTGATGCTTCCGAGGGCCGGCTC 2043
Db 1985 ATGATGCTGCATGCTCAACGGCTGCGCCCTGAGTCTGATGCTTCCGAGGGCCGGCTC 2044
Qy 2044 TGCTTCAAGATTGCGAGGAGCTCCGCTGTGAGGGAGATCCCAACCTCTCTGCGCCCT 2103
Db 2045 TGCTTCAAGATTGCGAGGAGCTCCGCTGTGAGGGAGATCCCAACCTCTCTGCGCCCT 2104
Qy 2104 CTCACAGCCCGAGGCTATCCAGAGGGGCTGAGGAAAGAGCCCATCCACGCGCTGTGCA 2163
Db 2105 CTCACAGCCCGAGGCTATCCAGAGGGGCTGAGGAAAGAGCCCATCCACGCGCTGTGCA 2164
Qy 2164 GCGGAGCTGGGAGGAAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT 2223
Db 2165 GCGGAGCTGGGAGGAAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT 2224

2224 TGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAAAATCAAGCCAATTTACCAACAG 2283
2225 TGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAAAATCAAGCCAATTTACCAACAG 2284
2284 ACCCTCCATGCCAGCCGAGAGAGCTTTGCGCAAGGGCCCCAGGGCCCCGCGCAGCTGAG 2343
2285 ACCCTCCATGCCAGCCGAGAGAGCTTTGCGCAAGGGCCCCAGGGCCCCGCGCAGCTGAG 2344
2344 GAGACAAACAGACAGAGCCCTTAAGCTCCAGGCTCTCTCCACAGAGCCGCCACAGAGCCA 2403
2345 GAGACAAACAGAGAGCCCTTAAGCTCCAGGCTCTCTCCACAGAGCCGCCACAGAGCCA 2404
2404 AACAAAGTCTCTCCCTTGACTTTGAGCAAGGAGAGTCTGGGATGTGGGAACCCCTTAACCT 2463
2405 AACAAAGTCTCTCCCTTGACTTTGAGCAAGGAGAGTCTGGGATGTGGGAACCCCTTAACCT 2464
2464 CTGTCTCTCCCTGGAGCCAGCCCTTGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2523
2465 CTGTCTCTCCCTGGAGCCAGCCCTTGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2524
2524 GTCCCGGAGCAGGAATCGACAGCTGGAAATAGAAATTTATTCCTCAACAGAGCTGTCCAG 2583
2525 GTCCCGGAGCAGGAATCGACAGCTGGAAATAGAAATTTATTCCTCAACAGAGCTGTCCAG 2584
2584 CCATTTTCTCTGGAGGAGCAGGACAAATTTCTCGTGCCCTCAGCATCGACAGCCTCTCC 2643
2585 CCATTTTCTCTGGAGGAGCAGGACAAATTTCTCGTGCCCTCAGCATCGACAGCCTCTCC 2644
2644 CTGTCCGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGCTCGCGGGACACCCCTG 2703
2645 CTGTCCGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGCTCGCGGGACACCCCTG 2704
2704 AGCTCAGGCTGACCTCTCGAGAGCCAGAGCCAGAGGCTCGAAGCTCCAGCTGGAACATG 2763
2705 AGCTCAGGCTGACCTCTCGAGAGCCAGAGCCAGAGGCTCGAAGCTCCAGCTGGAACATG 2764
2764 GTGCTGGCCCCGGGGCGGCCACCGACACACCCCAAGCTATTTCAATGTGTGAAAGTCCAA 2823
2765 GTGCTGGCCCCGGGGCGGCCACCGACACACCCCAAGCTATTTCAATGTGTGAAAGTCCAA 2824
2824 ATACAGTCTCTTAATGTGTGAAACACCTGACATCCGGGAGTTCCACCGGGTCAAAAGTGGGA 2883
2825 ATACAGTCTCTTAATGTGTGAAACACCTGACATCCGGGAGTTCCACCGGGTCAAAAGTGGGA 2884
2884 GACATCGCCATCGCATCAGCAGCCAGATCCAGCTCGAGCTTCAGCTTGGTCAACCAA 2943
2885 GACATCGCCATCGCATCAGCAGCCAGATCCAGCTCGAGCTTCAGCTTGGTCAACCAA 2944
2944 GACGGGCAAGCTGTTGCTAGCATGATGAGGTGCGAGCTCGGATCGAGCTTCAGCTGAGTGC 3003
2945 GACGGGCAAGCTGTTGCTAGCATGATGAGGTGCGAGCTCGGATCGAGCTTCAGCTGAGTGC 3004
3004 ACATGGCCCCGTGATGGAGCTTCCCTGGAGCTGGAGGTCAAGCATGGCCAGCTGGAG 3063
3005 ACATGGCCCCGTGATGGAGCTTCCCTGGAGCTGGAGGTCAAGCATGGCCAGCTGGAG 3064
3064 AACAGGCCCTTAACCTTGCCCTCCACCGCGGCTCCACCTGCGGAAAGAGCCCTTCTGT 3123
3065 AACAGGCCCTTAACCTTGCCCTCCACCGCGGCTCCACCTGCGGAAAGAGCCCTTCTGT 3124
3124 CTCGGTGACGATGCTGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATTTGCCAGGCC 3183
3125 CTCGGTGACGATGCTGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATTTGCCAGGCC 3184
3184 CCCCGGCTCAGATGGGAACCAAGGCTTCGACGACGAAGGTGGGGCAAGCAAGATGCC 3243
3185 CCCCGGCTCAGATGGGAACCAAGGCTTCGACGACGAAGGTGGGGCAAGCAAGATGCC 3244
3244 TCCAGGATTTTCAACCTGAGCCCTGCCCCACCTGCTGAAACCAACATCCGCCAGGTGAA 3303
3245 TCCAGGATTTTCAACCTGAGCCCTGCCCCACCTGCTGAAACCAACATCCGCCAGGTGAA 3304

3304 GAGACAGAGGAGATGGCAGGAGTTACTCTGGGGAACAAACAGGAGTCTTTTCTTGCC 3363
3305 GAGACAGAGGAGATGGCAGGAGTTACTCTGGGGAACAAACAGGAGTCTTTTCTTGCC 3364
3364 CCTGCTCAGTCCAGTTGGCTGACCCGCTTGGATCAGTGACCATTTGTTGGCAGACAGG 3423
3365 CCTGCTCAGTCCAGTTGGCTGACCCGCTTGGATCAGTGACCATTTGTTGGCAGACAGG 3424
3424 GGAGAGCAGCTTCCAGGCTTGGGTGAGAAAGGGTGGGCGAGCCCTTCGGCCCTCACCCTC 3483
3425 GGAGAGCAGCTTCCAGGCTTGGGTGAGAAAGGGTGGGCGAGCCCTTCGGCCCTCACCCTC 3484
3484 CAGGCTGCTGAGAGTGTCAAGTGTGTAAGGGCCCAAACTCAGGTTCAAGTGCAGAACCA 3543
3485 CAGGCTGCTGAGAGTGTCAAGTGTGTAAGGGCCCAAACTCAGGTTCAAGTGCAGAACCA 3544
3544 GGTGAGAGTATCCCGCGCTAGGTTAAAGGGGCGCTCTAAACCCCTTCGCTGGCCCTC 3603
3545 GGTGAGAGTATCCCGCGCTAGGTTAAAGGGGCGCTCTAAACCCCTTCGCTGGCCCTC 3604
3604 ACCTGGCAGCTCACCCCTTTTGGGTGTAGGGGAAAAGATGCTGACCTTGGGAAGGCT 3663
3605 ACCTGGCAGCTCACCCCTTTTGGGTGTAGGGGAAAAGATGCTGACCTTGGGAAGGCT 3664
3664 CCCTGGTAGAATACACACACTTTTCAAGTTGTTGCAACACAGGTCCTGAGTTGACCTCT 3724
3724 GGTTCAGCAAGGACCAAGAGGTGTAAAGTGAAGTGTCTCAGTCCCGACGACATGT 3783
3725 GGTTCAGCAAGGACCAAGAGGTGTAAAGTGAAGTGTCTCAGTCCCGACGACATGT 3784
3784 GCGCTTTGCTGCTGGCTACCACTTTTCCAGAGCAGCAGGCGCCCGAGCCCTTCAGGC 3843
3785 GCGCTTTGCTGCTGGCTACCACTTTTCCAGAGCAGCAGGCGCCCGAGCCCTTCAGGC 3844
3844 CCAGCATGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTTGAT 3903
3845 CCAGCATGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTTGAT 3904
3904 GCAGGATATGCTGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAGGAAAGCAG 3963
3905 GCAGGATATGCTGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAGGAAAGCAG 3964
3964 CAGAGAGAGCTCCGCGCCCCAGCCCACTATCAGTGTCCAGCTGCTGGTTCGCCAG 4023
3965 CAGAGAGAGCTCCGCGCCCCAGCCCACTATCAGTGTCCAGGCTGCTGGTTCGCCAG 4024
4024 AGCAGCTCAGCATCACCTGACACTCACCTTCCCTGCGCCCTGGCCAGAGGGTACTGC 4083
4025 AGCAGCTCAGCATCACCTGACACTCACCTTCCCTGCGCCCTGGCCAGAGGGTACTGC 4084
4084 CGAGGGCACTTTGCACTGTGATGACCTCAAGGCACTTTTTCATGGCTGCGCTTGGCAGGGC 4143
4085 CGAGGGCACTTTGCACTGTGATGACCTCAAGGCACTTTTTCATGGCTGCGCTTGGCAGGGC 4144
4144 AGGCGAGGGAGTGCACACTGTAGGAGCATAGCAAGCCAGAGAGATGGGGTGAAGGAGACA 4203
4145 AGGCGAGGGAGTGCACACTGTAGGAGCATAGCAAGCCAGAGAGATGGGGTGAAGGAGACA 4204
4204 GTCTTGAGCTGTCCACATGATGTGACTCTCTCAAACTCTTTCAGATTTTCTTAAGAATA 4263
4205 GTCTTGAGCTGTCCACATGATGTGACTCTCTCAAACTCTTTCAGATTTTCTTAAGAATA 4264
4264 GCACCCCTTCCCATTTGCCAGCTTACGCTCTTCCAGAGGAGGACTACTCAGGACTC 4323
4265 GCACCCCTTCCCATTTGCCAGCTTACGCTCTTCCAGAGGAGGACTACTCAGGACTC 4324
4324 ACGTAGCATTTAAATCAGCTGTGAATCGTCAGGGGGTGTCTGCTAGCCTCAACCTCTGGG 4383
4325 ACGTAGCATTTAAATCAGCTGTGAATCGTCAGGGGGTGTCTGCTAGCCTCAACCTCTGGG 4384
4384 GCAGGGGACGCCGAGACTCCGTGGGAGAGCTCATTTCCACATCTTTCACAGACAGGCTT 4443

Db 4385 GCAGGGGACCGCCGAGACTCCGTGGGAGAAGCTCAITCCACATCTTGCCAAACAGACGCTT 4444
Qy 4444 TGTCCAGCTGTCCACATTGAGTCAGACTGCTCCCGGGAGAGAGCCCGGGCCCCCAGCAC 4503
Db 4445 TGTCCAGCTGTCCACATTGAGTCAGACTGCTCCCGGGAGAGAGCCCGGGCCCCCAGCAC 4504
Qy 4504 ATAAAGAACTGCAGCTTGGTACTGCAGAGTCTGGGTGTGACAGAACTCTTTGTAAGCAA 4563
Db 4505 ATAAAGAACTGCAGCTTGGTACTGCAGAGTCTGGGTGTGACAGAACTCTTTGTAAGCAA 4564
Qy 4564 TAAAGTTGGGGTGATGACAAATGTTAAAAA 4595
Db 4565 TAAAGTTGGGGTGATGACAAATGTTAAAAA 4596

RESULT 2

US-10-172-118-975
; Sequence 975, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 975
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_003954
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-975

Query Match 99.6%; Score 4578.4; DB 17; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 5 GGGGACTGTGCCGTGTGGAACGTGTAGCTGTTGA-AGGTGACTCTGTTACATTGAGG 63
Db 5 GGGGACTGTGCCGTGTGGAACGTGTAGCTGTTGAGAGTGTGACTCTGTTACATTGAGG 64
Qy 64 ATGTTTGGAGTAGTATGTGTGGCAGAGGCACACATAAAGCAGAGACCCCTTTGCC 123
Db 65 ATGTTTGGAGTAGTATGTGTGGCAGAGGCACACATAAAGCAGAGACCCCTTTGCC 124
Qy 124 CTTGCTTTCTCCCAACCCAGGCTGACCTGTGTTCTCCAGGTCTGGGATTTAAGT 183
Db 125 CTTGCTTTCTCCCAACCCAGGCTGACCTGTGTTCTCCAGGTCTGGGATTTAAGT 184
Qy 184 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAAGCTGGGAGATGGCAGTGATG 243
Db 185 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAAGCTGGGAGATGGCAGTGATG 244
Qy 244 GAAATGGCTGCCAGGTGCCCTCGCTCAGAGTGGGACAGAGAACTCCCCAAAG 303
Db 245 GAAATGGCTGCCAGGTGCCCTCGCTCAGAGTGGGACAGAGAACTCCCCAAAG 304
Qy 304 CCAAGGAGAGACGCGCCACTGGGGAAGAAACAGAGCTCCGTCTACNAGTTGAGGCC 363
Db 305 CCAAGGAGAGACGCGCCACTGGGGAAGAAACAGAGCTCCGTCTACNAGTTGAGGCC 364

Qy 364 GTGAGAAAGAGCCCTGTGTCTTGGGAAAGTGGAGATCTCTGAATGACGTGATTACCAAG 423
Db 365 GTGAGAAAGAGCCCTGTGTCTTGGGAAAGTGGAGATCTCTGAATGACGTGATTACCAAG 424
Qy 424 GGCACAGCAAGGAAGGCTCCGAGGACGGCCAGCTGCGCATCTCTATCATGCCCCAGGCT 483
Db 425 GGCACAGCAAGGAAGGCTCCGAGGACGGCCAGCTGCGCATCTCTATCATGCCCCAGGCT 484
Qy 484 GAGTGTGAGAAATAGCCAGAGTTTCAGCCCCACCTTTTTCAGAACGCAATTTTCATCGCTGG 543
Db 485 GAGTGTGAGAAATAGCCAGAGTTTCAGCCCCACCTTTTTCAGAACGCAATTTTCATCGCTGG 544
Qy 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCCATGCT 603
Db 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCCATGCT 604
Qy 604 ACAGAGGCAAAATGGCCCGTGTGTGTTGGAAGGAAAGCGTTCGAGCAAAAGCCCCGGAAG 663
Db 605 ACAGAGGCAAAATGGCCCGTGTGTGTTGGAAGGAAAGCGTTCGAGCAAAAGCCCCGGAAG 664
Qy 664 AAAAGGAGAAAGAGAGCTCAAAAGTCCCTGCTCATGACAGAGTGGCCCTTGGCCAAAACC 723
Db 665 AAAAGGAGAAAGAGAGCTCAAAAGTCCCTGCTCATGACAGAGTGGCCCTTGGCCAAAACC 724
Qy 724 CTCCCCAGGACCCCTGAGCAGAGAGTGCACCATCCAGTCGAGGAGATGAGTCTCCA 783
Db 725 CTCCCCAGGACCCCTGAGCAGAGAGTGCACCATCCAGTCGAGGAGATGAGTCTCCA 784
Qy 784 CTCGGCGCCCATATGTTTAGAAACACCCCGCAGTTTCAACAGCCCTCTGAAAGAAACGAGC 843
Db 785 CTCGGCGCCCATATGTTTAGAAACACCCCGCAGTTTCAACAGCCCTCTGAAAGAAACGAGC 844
Qy 844 CTGCGCAACTCTGTTTAAAGCAGCTTGGCAGAGGCTTACGCGCGGCTCTGCTCGATCA 903
Db 845 CTGCGCAACTCTGTTTAAAGCAGCTTGGCAGAGGCTTACGCGCGGCTCTGCTCGATCA 904
Qy 904 GAACTCCACAAACTGATCAGGCCCTTGCAATGTCTGAAACACACGCTGTGAAACTGCACCAC 963
Db 905 GAACTCCACAAACTGATCAGGCCCTTGCAATGTCTGAAACACACGCTGTGAAACTGCACCAC 964
Qy 964 CCCAGGACGAGGCGCCCTGCCCCGTGCCACGACGCCCTTCCCTTATAGAGACTGCTT 1023
Db 965 CCCAGGACGAGGCGCCCTGCCCCGTGCCACGACGCCCTTCCCTTATAGAGACTGCTT 1024
Qy 1024 CATCCCTTCCCATTCACCCCTTCAGCCCTTGGAAACCTCACCCCTCTGGAGTCTTCTG 1083
Db 1025 CATCCCTTCCCATTCACCCCTTCAGCCCTTGGAAACCTCACCCCTCTGGAGTCTTCTG 1084
Qy 1084 GGCAACTGGCTGTGTAGACAGCCAGAAACCCCTTGGCTGACCCACACCTGAGCAAACTG 1143
Db 1085 GGCAACTGGCTGTGTAGACAGCCAGAAACCCCTTGGCTGACCCACACCTGAGCAAACTG 1144
Qy 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGTCT 1203
Db 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGTCT 1204
Qy 1204 CGTGTGCCATGAGAGTTTCTGTGAGAGAAATACCTAGTGCATGCTCTGCAAGCGAGC 1263
Db 1205 CGTGTGCCATGAGAGTTTCTGTGAGAGAAATACCTAGTGCATGCTCTGCAAGCGAGC 1264
Qy 1264 GTGAGCTCAAGCCAGGCCCCACAGCTGACCCCTGGCCAAAGACCTGGGACGACGCGG 1323
Db 1265 GTGAGCTCAAGCCAGGCCCCACAGCTGACCCCTGGCCAAAGACCTGGGACGACGCGG 1324
Qy 1324 TCCAGATCCCGGAGCCCGAGCCCAAAACTGAGGACAAAGAGGTGTCTGTCTACTGAG 1383
Db 1325 TCCAGATCCCGGAGCCCGAGCCCAAAACTGAGGACAAAGAGGTGTCTGTCTACTGAG 1384
Qy 1384 AAATCTCAAGCAGTGGATTATGAGTACCGAGAGAAAGTCCACTGGGCGCACGACGCTC 1443
Db 1385 AAATCTCAAGCAGTGGATTATGAGTACCGAGAGAAAGTCCACTGGGCGCACGACGCTC 1444
Qy 1444 CGCTGGGACAGGCTCTCTCGGAGAGGTGCACAGGATGGAGGACAAAGCAGACTGGCTTC 1503

Db 1445 |||||CGCTGGGCGAGAGGCTCTCTCGGAGAGGTGCACAGGATGGAGGACAAGCAGACTGGCTTC 1504
Qy |||||CAGTGGCGCTGTCAAAAAGGTCGGCTCGGAAGTATTTGGGGCAGAGGAGCTGATGGCATGT 1563
Db 1505 CAGTGGCGCTGTCAAAAAGGTCGGCTCGGAAGTATTTGGGGCAGAGGAGCTGATGGCATGT 1564
Qy 1564 GCAGGATGACCTCACCCAGAAATTTGTCCTTTGATGGAGCTGTGAGAGAGGGCCCTTGG 1623
Db 1565 GCAGGATGACCTCACCCAGAAATTTGTCCTTTGATGGAGCTGTGAGAGAGGGCCCTTGG 1624
Qy 1624 GTCAACATCTTCATGGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAG 1683
Db 1625 GTCAACATCTTCATGGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAG 1684
Qy 1684 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGAGCCCTGGAGGGTCTCGAA 1743
Db 1685 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGAGCCCTGGAGGGTCTCGAA 1744
Qy 1744 TACTCCACTCACAAAGGATTTCTCATGGGACGTCAAAGCTGACAACTGCTCTCTGTCC 1803
Db 1745 TACTCCACTCAAGAGGATTTCTCATGGGACGTCAAAGCTGACAACTGCTCTCTGTCC 1804
Qy 1804 AGCGATGGAGCCACGAGCCCTCTGTGACTTTGGCCATGTGTGTCTTTCAAACCTGAT 1863
Db 1805 AGCGATGGAGCCACGAGCCCTCTGTGACTTTGGCCATGTGTGTCTTTCAAACCTGAT 1864
Qy 1864 GGCTGTGGGAAAGTCTTGTCTCACAGGGGACTACATCCTTGGCACAGAGCCACATGGCT 1923
Db 1865 GGCTGTGGGAAAGTCTTGTCTCACAGGGGACTACATCCTTGGCACAGAGCCACATGGCT 1924
Qy 1924 CCGAGGTGTGTCTGGGAGGAGCTGGGACGCCAAGGTGATGTCTGGAGCAGCTGTGT 1983
Db 1925 CCGAGGTGTGTCTGGGAGGAGCTGGGACGCCAAGGTGATGTCTGGAGCAGCTGTGT 1984
Qy 1984 ATGATGTGCACATGTCTCAAGGCTGCGCACCCCTTG3ACTCAGTTCTTCGAGGGCGCGTC 2043
Db 1985 ATGATGTGCACATGTCTCAAGGCTGCGCACCCCTTG3ACTCAGTTCTTCGAGGGCGCGTC 2044
Qy 2044 TGGCTCAAGATTGCCAGGAGCCCTCGCCTGTGAGGGAGATCCCAACCTCTCTGGCCCT 2103
Db 2045 TGGCTCAAGATTGCCAGGAGCCCTCGCCTGTGAGGGAGATCCCAACCTCTCTGGCCCT 2104
Qy 2104 CTCACAGCCAGGCCATCCAGAGGGCTGAGGAAAGAGCCATCCACCGGCTGTCTGCA 2163
Db 2105 CTCACAGCCAGGCCATCCAGAGGGCTGAGGAAAGAGCCATCCACCGGCTGTCTGCA 2164
Qy 2164 GCGAGCTGGAGGGAAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGAGCCCT 2223
Db 2165 GCGAGCTGGAGGGAAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGAGCCCT 2224
Qy 2224 TGGAGGGAGAAATATAAGAACCAAGACATCCACCGCCCAATCAAGCCNATTACCAACAG 2283
Db 2225 TGGAGGGAGAAATATAAGAACCAAGACATCCACCGCCCAATCAAGCCNATTACCAACAG 2284
Qy 2284 ACCCTCATGCCAGCGAGAGGCTTTGGCCAAAGGGCCCGAGGGCCCGGCCAGCTGAG 2343
Db 2285 ACCCTCATGCCAGCGAGAGGCTTTGGCCAAAGGGCCCGAGGGCCCGGCCAGCTGAG 2344
Qy 2344 GAGACAAAGGAGAGCCCTTAAGCTCCAGCTCTCTCTCCACAGAGCCCGCCAGAGCCCA 2403
Db 2345 GAGACAAAGGAGAGCCCTTAAGCTCCAGCTCTCTCTCCACAGAGCCCGCCAGAGCCCA 2404
Qy 2404 AACAGTCTCTCCCTTGTACTTTGAGCAAGAGAGGTCTGGGATGTGGGAAACCCCTTACCT 2463
Db 2405 AACAGTCTCTCCCTTGTACTTTGAGCAAGAGAGGTCTGGGATGTGGGAAACCCCTTACCT 2464
Qy 2464 CTGTCTCTCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2523
Db 2465 CTGTCTCTCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2524
Qy 2524 GTCCCGGAGCAGGAATGACAGCTGGAAATAGAAATATTTCCTCAACAGGCTCTCCAG 2583

Db 2525 GTCCCGGAGCGGAACTGACAGAGCTGGAATATTAATTTCTCAACAGGCTGTCCAG 2584
Qy 2584 CCATTTTCTCTGGAGGACGAGCAAAATTTCTCTGTCCTCAGCATCGACAGCTCTCC 2643
Db 2585 CCATTTTCTCTGGAGGACGAGCAAAATTTCTCTGTCCTCAGCATCGACAGCTCTCC 2644
Qy 2644 CTGTCCGATCAGAGTGAAGAACCCTCAAGAGCCCTCTCAAGCTCGCGGACACCTG 2703
Db 2645 CTGTCCGATCAGAGTGAAGAACCCTCAAGAGCCCTCTCAAGCTCGCGGACACCTG 2704
Qy 2704 AGCTCAGCGCTACACTCTCGAGCAGCAGCCGAGGCTCGAAGCTCAGCTGGAACATG 2763
Db 2705 AGCTCAGCGCTACACTCTCGAGCAGCAGCCGAGGCTCGAAGCTCAGCTGGAACATG 2764
Qy 2764 GTCTGSCCCGCGGGCGGCCACCGACACCCCAAGCTATTTCATATGTGTGAAGTCCAA 2823
Db 2765 GTCTGSCCCGCGGGGGGCCACCGACACCCCAAGCTATTTCATATGTGTGAAGTCCAA 2824
Qy 2824 ATACAGTCTCTTAATGTGTGAACACTGTGCATATCGGGAGTTCCACCGGCTCAAGTGGGA 2883
Db 2825 ATACAGTCTCTTAATGTGTGAACACTGTGCATATCGGGAGTTCCACCGGCTCAAGTGGGA 2884
Qy 2884 GACATCGCCACTGGCATCAGCAGCCAGATCCCAAGCTTTTCAATGTGTGAAGTCCAA 2943
Db 2885 GACATCGCCACTGGCATCAGCAGCCAGATCCCAAGCTTTTCAATGTGTGAAGTCCAA 2944
Qy 2944 GAGGGCAGCTGTTCGTCTACGACATGGAGGTGCAGACTCGGGCATCGACCTCGAGTGC 3003
Db 2945 GAGGGCAGCTGTTCGTCTACGACATGGAGGTGCAGACTCGGGCATCGACCTCGAGTGC 3004
Qy 3004 ACATGSCCCTGTATGTCAGCTTGGCTGAGCTGAGGGTCAAGCATGAGCACTGGAG 3063
Db 3005 ACATGSCCCTGTATGTCAGCTTGGCTGAGCTGAGGGTCAAGCATGAGCACTGGAG 3064
Qy 3064 AACAGGCCCTTAAACCTGSCCTCCACCGCCGCTCCACACTGCGGAAAGAGCCTTCTG 3123
Db 3065 AACAGGCCCTTAAACCTGSCCTCCACCGCCGCTCCACACTGCGGAAAGAGCCTTCTG 3124
Qy 3124 CTGCTGACAGTGTCTCCCTGAAAAACACAGGCTCAGCGCTTCCAGGGGATTGCCAGCC 3183
Db 3125 CTGCTGACAGTGTCTCCCTGAAAAACACAGGCTCAGCGCTTCCAGGGGATTGCCAGCC 3184
Qy 3184 CCGCGGCTCAGTGGGAAACAGGGCTTCGACGACAGCAAGTGGGGCAAGCAATGTC 3243
Db 3185 CCGCGGCTCAGTGGGAAACAGGGCTTCGACGACAGCAAGTGGGGCAAGCAATGTC 3244
Qy 3244 TCCAGGATTTACACCTGAGCCCTGCCACCTGCTGAAAAAACAATCCGCCACGTGAA 3303
Db 3245 TCCAGGATTTACACCTGAGCCCTGCCACCTGCTGAAAAAACAATCCGCCACGTGAA 3304
Qy 3304 GAGACAGGAGGATGGCAGGAGTTACCTGGGAAAAACAAAACAGGATCTTTTCTGTC 3363
Db 3305 GAGACAGGAGGATGGCAGGAGTTACCTGGGAAAAACAAAACAGGATCTTTTCTGTC 3364
Qy 3364 CTTGCTCCAGTTGGCTGACCCGCTTGGATCAGTGACCAATTTTGTGGCAGACAGG 3423
Db 3365 CTTGCTCCAGTTGGCTGACCCGCTTGGATCAGTGACCAATTTTGTGGCAGACAGG 3424
Qy 3424 GAGAGCAGCTTCAGAGCTGGGTCAAGAGGGTGGGAGCCCTTCGGCCCTCACCCCTC 3483
Db 3425 GAGAGCAGCTTCAGAGCTGGGTCAAGAGGGTGGGAGCCCTTCGGCCCTCACCCCTC 3484
Qy 3484 CAGGCTGCTGAGAGTCAAGTGTGTAAAGGGCCCAAACTCAGGTTCAAGTGCAGAACCA 3543
Db 3485 CAGGCTGCTGAGAGTCAAGTGTGTAAAGGGCCCAAACTCAGGTTCAAGTGCAGAACCA 3544
Qy 3544 GGTCAAGAGTATGCCCGCCGTAGGTTAAGGGGGCCCTCTAAACCCCTTCCTGGGCTC 3603
Db 3545 GGTCAAGAGTATGCCCGCCGTAGGTTAAGGGGGCCCTCTAAACCCCTTCCTGGGCTC 3604
Qy 3604 ACCTGGCAGCTCACCCCTTTGGGTGTAGGGGAAAAAGATGCTGACCTCGGAAAGGCT 3663
Db 3605 ACCTGGCAGCTCACCCCTTTGGGTGTAGGGGAAAAAGATGCTGACCTCGGAAAGGCT 3664

3664 CCCTGCTAGAAATACACCACTTTTCAGGTGTTGTCACACAGGTCCTGAGTTGACCTCT 3723
3665 CCCTGGTAGAATACACCACTTTTCAGGTGTTGTCACACAGGTCCTGAGTTGACCTCT 3724
3724 GCTTCAGCCAGGACCAAGAGGTGTGTAAAGTGAAGTGGTTCTCAGTCCCCAGACATGT 3783
3725 GCTTCAGCCAGGACCAAGAGGTGTGTAAAGTGAAGTGGTTCTCAGTCCCCAGACATGT 3784
3784 GCCCTTTGCTGCTGCTACCACTCTTCCCAAGAGAGCAGGCCCCGAGGCCCTTTTCAGGC 3843
3785 GCCCTTTGCTGCTGCTACCACTCTTCCCAAGAGAGCAGGCCCCGAGGCCCTTTTCAGGC 3844
3844 CCAGCACTGCCCCAGACTGCTGGGACATCAGTTCCTCATCTGTAAGGTGAAGGTGAT 3903
3845 CCAGCACTGCCCCAGACTGCTGGGACATCAGTTCCTCATCTGTAAGGTGAAGGTGAT 3904
3904 CGAGGATATGCTGTGACAGAAACAGTCTGTGGATGGACATGATCAGTGTCTAAGGAAAGCAG 3963
3905 GCAGGATATGCTGTGACAGAAACAGTCTGTGGATGGACATGATCAGTGTCTAAGGAAAGCAG 3964
3964 CAGAGAGAGACGTCCGGGGCCCCAGCCCCCACTATCAGTGTCCAGCGTGTGTTCCCGAG 4023
3965 CAGAGAGAGACGTCCGGGGCCCCAGCCCCCACTATCAGTGTCCAGCGTGTGTTCCCGAG 4024
4024 AGCAGACTCAGCATCAGCTGACACTCACCTGCCCCCTGCCCTGGCCAGAGGTACTGC 4083
4025 AGCAGACTCAGCATCAGCTGACACTCACCTGCCCCCTGCCCTGGCCAGAGGTACTGC 4084
4084 CCACGGCACTTTGCACTCTGTATGACCTCAAGACACTTTTCATGGCTGCCCTCTGGCAGGCG 4143
4085 CACGGCACTTTGCACTCTGTATGACCTCAAGACACTTTTCATGGCTGCCCTCTGGCAGGCG 4144
4144 AGGGCAGGGCAGTGAACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGACACA 4203
4145 AGGGCAGGGCAGTGAACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGACACA 4204
4204 GTCTTGAGCTGTCCACATGATGTGACTCTCAAACTCTTCCAGATTTCTTAAGAATA 4263
4205 GTCTTGAGCTGTCCACATGATGTGACTCTCAAACTCTTCCAGATTTCTTAAGAATA 4264
4264 GCACCCCTTTCCCACTTGCCAGCTTAGCCCTTTCTCCAGGGGAGCTACTCAGGACTC 4323
4265 GCACCCCTTTCCCACTTGCCAGCTTAGCCCTTTCTCCAGGGGAGCTACTCAGGACTC 4324
4324 ACGTAGCATTAATCAGCTGTGAATCGTCAAGGGGTGTCTGCTAGCCTCAACCTCCTGGG 4383
4325 ACGTAGCATTAATCAGCTGTGAATCGTCAAGGGGTGTCTGCTAGCCTCAACCTCCTGGG 4384
4384 GCAGGGGACGCGAGACTCCGTGGGAGAGCTCATTTCCACATCTTGCCAGACAGCCTT 4443
4385 GCAGGGGACGCGAGACTCCGTGGGAGAGCTCATTTCCACATCTTGCCAGACAGCCTT 4444
4444 TGTCAGCTGTCCACATTCAGTTCAGTCTCTCCGGGAGAGAGCCCGGCCCCAGCAC 4503
4445 TGTCAGCTGTCCACATTCAGTTCAGTCTCTCCGGGAGAGAGCCCGGCCCCAGCAC 4504
4504 ATAAAGAACTGCAGCCTTGGTACTGCGAGAGTCTGGGTTGTAGAGAACTCTTTTGAAGCAA 4563
4505 ATAAAGAACTGCAGCCTTGGTACTGCGAGAGTCTGGGTTGTAGAGAACTCTTTTGAAGCAA 4564
4564 TAAAGTTGGGGTGATGACAAATGTTAAAAA 4595
4565 TAAAGTTGGGGTGATGACAAATGTTAAAAA 4596

RESULT 3
US-10-342-887-975
; Sequence 975, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong

; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 975
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-342-887-975

Query Match 99.6%; Score 4578.4; DB 18; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 5 GGGGACTGTGCCGTGTGAAACGTGTAGCTGTTGA-AGGTGACTCTGTTACCATGAGG 63
DB 5 GGGGACTGTGCCGTGTGAAACGTGTAGCTGTTGAAGAGGTGACTCTGTTACCATGAGG 64
QY 64 ATGTTTGAGATGATGTGTGGCAGAGCACATATAACAGCAGACACCTTTGCC 123
DB 65 ATGTTTGAGATGATGTGTGGCAGAGCACATATAACAGCAGACACCTTTGCC 124
QY 124 CTTGCTTTCTCCCCAACCAAGGCTCACCTGTGTTCTCCAGGCTCTGGGATTTCTAAGT 183
DB 125 CTTGCTTTCTCCCCAACCAAGGCTGACCTGTGTTCTCCAGGCTCTGGGATTTCTAAGT 184
QY 184 GACCTGCTCTGTGTTGTTGTTCTCTCAGGATGAGCACAAAGCTTGGGAGATGGCAGTGATG 243
DB 185 GACCTGCTCTGTGTTGTTGTTCTCTCAGGATGAGCACAAAGCTTGGGAGATGGCAGTGATG 244
QY 244 GAAATGGCTTCCCGAGGTGCCCCCTGGCTCAGCAGTGGGCGAGCAAGAACTCCCCAAG 303
DB 245 GAAATGGCTTCCCGAGGTGCCCCCTGGCTCAGCAGTGGGCGAGCAAGAACTCCCCAAG 304
QY 304 CCAAGGAGAGAGACGCCGCCACTTGGGGAAGAAACAGAGCTCCGTCTACAAGCTTGAGGCC 363
DB 305 CCAAGGAGAGAGACGCCGCCACTTGGGGAAGAAACAGAGCTCCGTCTACAAGCTTGAGGCC 364
QY 364 GTGGAGAAAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCTCTGAATGAGCGTGAATACCAAG 423
DB 365 GTGGAGAAAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCTCTGAATGAGCGTGAATACCAAG 424
QY 424 GGCACAGCAAGAGAGGCTCCGAGGCGAGGCTGCACTCTCTATCATCGCCAGGCT 483
DB 425 GGCACAGCAAGAGAGGCTCCGAGGCGAGGCTGCACTCTCTATCATCGCCAGGCT 484
QY 484 GAGTGTGAGAAATAGCAAGAGTTTCAAGCCCACTTTTTCAGAACGCAATTTTTCATCGCTGGG 543
DB 485 GAGTGTGAGAAATAGCAAGAGTTTCAAGCCCACTTTTTCAGAACGCAATTTTTCATCGCTGGG 544
QY 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAATGTGGCCCATGCT 603
DB 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAATGTGGCCCATGCT 604
QY 604 ACAGAGGCAAAATGGCCCGTGTGTTGGAGGAAAGCGTCCGAGCAAAAGCCCGGAG 663
DB 605 ACAGAGGCAAAATGGCCCGTGTGTTGGAGGAAAGCGTCCGAGCAAAAGCCCGGAG 664
QY 664 AAAACGAAAGAGAGAGTCAAAAGTCCCTGGCTCATGCGAGGAGTGGCCCTTTGGCCAAACCC 723

Db 665 AAACGGAGAGAGAGAGCTCAAGTCCCTGGCTCATGACAGAGTGGCTTTGGCCAAACCC 724
Qy 724 CTCCCAGAGCCCTTGACAGAGAGCTGACCATCCAGTCCAGTGACAGAGATGAGTCTCCA 783
Db 725 CTCCCAGAGCCCTTGACAGAGAGCTGACCATCCAGTCCAGTGACAGAGATGAGTCTCCA 784
Qy 784 CTGGGCGCCCATATGTTAGAAAAACCCCGCAGTTTACCAAGCCCTCTGAAAGAAACCAAGGC 843
Db 785 CTGGGCGCCCATATGTTAGAAAAACCCCGCAGTTTACCAAGCCCTCTGAAAGAAACCAAGGC 844
Qy 844 CTGGGCAACTCTGTTTAAAGCAGCTTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGTCA 903
Db 845 CTGGGCAACTCTGTTTAAAGCAGCTTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGTCA 904
Qy 904 GAACTCCACAACTGATCAGAGCCCTTGCAATGTCTGAAACACAGTGTGAAACTGCACAC 963
Db 905 GAACTCCACAACTGATCAGAGCCCTTGCAATGTCTGAAACACAGTGTGAAACTGCACAC 964
Qy 964 CCCAGGACGAGGCCCCCTGCCCCGCTGCCACGCAACCCCTTCCCTATPAGCAGACTGCCT 1023
Db 965 CCCAGGACGAGGCCCCCTGCCCCGCTGCCACGCAACCCCTTCCCTATPAGCAGACTGCCT 1024
Qy 1024 CATCCCTTCCCATTTCCACCCCTCTCCAGCCCTGGAACCTCACCCTCTGGAGTCTTCTCTG 1083
Db 1025 CATCCCTTCCCATTTCCACCCCTCTCCAGCCCTGGAACCTCACCCTCTGGAGTCTTCTCTG 1084
Qy 1084 GGCAAACTGGGCTGTGTAGACAGCAGAAACCTTGGCTGACCCACACCTCAGCAAACTG 1143
Db 1085 GGCAAACTGGGCTGTGTAGACAGCAGAAACCTTGGCTGACCCACACCTGAGCAAACTG 1144
Qy 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCCACACTGGAGCCGAGCTGCTGTCT 1203
Db 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCCACACTGGAGCCGAGCTGCTGTCT 1204
Qy 1204 CGTGGTCCCATGAGAAGTTTCTGTGAGAGAACTCTAGTGTGATGTCTGCAAGGAGC 1263
Db 1205 CGTGGTCCCATGAGAAGTTTCTGTGAGAGAACTCTAGTGTGATGTCTGCAAGGAGC 1264
Qy 1264 GTGAGCTCAAGCCAGGCCCCAGAGCTGACAGCTGGCCAGAGCTGGGAGCAGCGGGC 1323
Db 1265 GTGAGCTCAAGCCAGGCCCCAGAGCTGACAGCTGGCCAGAGCTGGGAGCAGCGGGC 1324
Qy 1324 TCCAGATCCCGGAGGCCAGCCCCAAACTGAGGACCAACAGAGGTGCTCTCACTGAG 1383
Db 1325 TCCAGATCCCGGAGGCCAGCCCCAAACTGAGGACCAACAGAGGTGCTCTCACTGAG 1384
Qy 1384 AAATCAAGCCAGTGGATTTAGTACCGAGAGAGTTCCTGGGCCACGCAACAGCTC 1443
Db 1385 AAATCAAGCCAGTGGATTTAGTACCGAGAGAGTTCCTGGGCCACGCAACAGCTC 1444
Qy 1444 CGCTGGGCGAGGCTCTTGGAGAGGTGTCACAGGATGGAGGACAGCAGCTGGCTTC 1503
Db 1445 CGCTGGGCGAGGCTCTTGGAGAGGTGTCACAGGATGGAGGACAGCAGCTGGCTTC 1504
Qy 1504 CAGTGGCTGTCAAAAGGCTCGGCTGGAAGTATTTGGGCGAGGAGCTGATGGCATGT 1563
Db 1505 CAGTGGCTGTCAAAAGGCTCGGCTGGAAGTATTTGGGCGAGGAGCTGATGGCATGT 1564
Qy 1564 GCAGGATGACCTCACCCAGAAATGCTCCCTTGTATGGAGCTGTGAGAGAGGCGCTTGG 1623
Db 1565 GCAGGATGACCTCACCCAGAAATGCTCCCTTGTATGGAGCTGTGAGAGAGGCGCTTGG 1624
Qy 1624 GTCAACATCTTCAGGAGCTCTGGAAGGTGGCTCCCTGGGCCAGCTGGTCAAGAGCAG 1683
Db 1625 GTCAACATCTTCAGGAGCTCTGGAAGGTGGCTCCCTGGGCCAGCTGGTCAAGAGCAG 1684
Qy 1684 GGCTGTCTCCAGAGGACCGGCGCTGTACTACCTGGGCGAGGCGCTGGAGGGTCTGAA 1743
Db 1685 GGCTGTCTCCAGAGGACCGGCGCTGTACTACCTGGGCGAGGCGCTGGAGGGTCTGAA 1744
Qy 1744 TACCTCACTCAAGAGGATTTCTGATGGGAGCTCAAGCTGCAAGCTGCTCTCTGTC 1803
Db 1745 TACCTCACTCAAGAGGATTTCTGATGGGAGCTCAAGCTGCAAGCTGCTCTCTGTC 1804

Qy 1804 AGCGATGGAGCCACGACGCCCTCTGTGAATTTGGCCATGCTGTGTCTTTCAACCTGAT 1863
Db 1805 AGCGATGGAGCCACGACGCCCTCTGTGTGATTTTGGCCATGCTGTGTCTTTCAACCTGAT 1864
Qy 1864 GGCTGGGAAAGTCTTTGCTCAAGGGGACTACATCTCTGGCACAGAGACCAATGGCT 1923
Db 1865 GGCTGGGAAAGTCTTTGCTCAAGGGGACTACATCTCTGGCACAGAGACCAATGGCT 1924
Qy 1924 CCGGAGTGTGTGGGCGAGGAGCTGACGCCAAAGTGTGATGTCTGGAGCAGCTGTGT 1983
Db 1925 CCGGAGTGTGTGGGCGAGGAGCTGACGCCAAAGTGTGATGTCTGGAGCAGCTGTGT 1984
Qy 1984 ATGATGTGCACATGCTCAACGGTGCACGCCCTTGGACTCAGTTCTTCCAGGGGCGCTC 2043
Db 1985 ATGATGTGCACATGCTCAACGGTGCACGCCCTTGGACTCAGTTCTTCCAGGGGCGCTC 2044
Qy 2044 TGCCTCAAGATTGTCAGGAGCTTCCGCTGTGTAGGAGATGCCACCTCTCTGGGCGCT 2103
Db 2045 TGCCTCAAGATTGTCAGGAGCTTCCGCTGTGTAGGAGATGCCACCTCTCTGGGCGCT 2104
Qy 2104 CTACAGCCAGGCGCATCCAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGCA 2163
Db 2105 CTACAGCCAGGCGCATCCAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGCA 2164
Qy 2164 GCGGAGCTGGAGGAAAGGTGAACCGGCACTACAGCAAGTGGAGGTCTCAAGAGCCCT 2223
Db 2165 GCGGAGCTGGAGGAAAGGTGAACCGGCACTACAGCAAGTGGAGGTCTCAAGAGCCCT 2224
Qy 2224 TGAAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAATCAAGGCAATTAACACAG 2283
Db 2225 TGAAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAATCAAGGCAATTAACACAG 2284
Qy 2284 ACCCTCATCCAGCCGAGAGAGCTTTCCGCAAGGGCCCCAGGGCCCCGCGCCAGCTGAG 2343
Db 2285 ACCCTCATCCAGCCGAGAGAGCTTTCCGCAAGGGCCCCAGGGCCCCGCGCCAGCTGAG 2344
Qy 2344 GAGCAACAGCAGAGCCCTTAAGCTCAGCTCTCTCCAGCAGAGCCGCCAGAGGCA 2403
Db 2345 GAGCAACAGCAGAGCCCTTAAGCTCAGCTCTCTCCAGCAGAGCCGCCAGAGGCA 2404
Qy 2404 AACAGTCTCTCCCTTGAATTTGAGCAAGGAGGTCTGGGATGTGGGAAACCTTACCT 2463
Db 2405 AACAGTCTCTCCCTTGAATTTGAGCAAGGAGGTCTGGGATGTGGGAAACCTTACCT 2464
Qy 2464 CTGTCTCTCTGAGCCAGCCCTGCGCAGAAACCCAGCTCACCAGAGCGGAAAGCAAC 2523
Db 2465 CTGTCTCTCTGAGCCAGCCCTGCGCAGAAACCCAGCTCACCAGAGCGGAAAGCAAC 2524
Qy 2524 GTCCCGGAGCAGGAACTGACAGAGCTGGAATAGAAATTTCTCTCAACAGCTGTCCAG 2583
Db 2525 GTCCCGGAGCAGGAACTGACAGAGCTGGAATAGAAATTTCTCTCAACAGCTGTCCAG 2584
Qy 2584 CCATTTTCTCTGGAGGAGCAGGAGCAATTTCTCTGCTGCTCAGCATCGACAGCTCTCC 2643
Db 2585 CCATTTTCTCTGGAGGAGCAGGAGCAATTTCTCTGCTGCTCAGCATCGACAGCTCTCC 2644
Qy 2644 CTGTGAGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGCTCGCGGAGCACCTGT 2703
Db 2645 CTGTGAGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGCTCGCGGAGCACCTGT 2704
Qy 2704 AGCTCAGGCGTACACTCTCGAGCAGCAGGCGGAGCTCGAAGCTCAGCTGGAACATG 2763
Db 2705 AGCTCAGGCGTACACTCTCGAGCAGCAGGCGGAGCTCGAAGCTCAGCTGGAACATG 2764
Qy 2764 GTGTGCGCGCGGCGGCGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAGTCCAA 2823
Db 2765 GTGTGCGCGCGGCGGCGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAGTCCAA 2824
Qy 2824 ATACAGTCTCTTAATGAGTGAACCTTGCAATCGGAGGTTCACCCGGGTCAAGAGTGGGA 2883
Db 2825 ATACAGTCTCTTAATGAGTGAACCTTGCAATCGGAGGTTCACCCGGGTCAAGAGTGGGA 2884

QY	2884	GACATCCCACTGGCATCAGAGCAGATCCCAGCTGCAGCCTTCAGCTTGGTCA	2944
DB	2885	GACATCCCACTGGCATCAGAGCAGATCCCAGCTGCAGCCTTCAGCTTGGTCA	2944
QY	2944	GAGGGCAGCCTGTTTCGCTACGACATGAGAGTGCACAGCTCGGGCATCGAC	3003
DB	2945	GAGGGCAGCCTGTTTCGCTACGACATGAGAGTGCACAGCTCGGGCATCGAC	3004
QY	3004	ACACTGGCCCTGATGTGCAGCTTCGCTCGAGCTGGAGGCTCAAGCATGCCAG	3063
DB	3005	ACACTGGCCCTGATGTGCAGCTTCGCTCGAGCTGGAGGCTCAAGCATGCCAG	3064
QY	3064	AACAGGCCCTAAACCTCGCTCCACCGCCGGCTCCACACTGCCGAAAGACAG	3123
DB	3065	AACAGGCCCTAAACCTCGCTCCACCGCCGGCTCCACACTGCCGAAAGACAG	3124
QY	3124	CTCGGTCCAGATGCTGCTGAAACACAGGCTCAGCCGTTCCAGGGGATTCAG	3183
DB	3125	CTCGGTCCAGATGCTGCTGAAACACAGGCTCAGCCGTTCCAGGGGATTCAG	3184
QY	3184	CCCCGGCTCACAGTGGGAAACAGGGCTTCGACAGCAGCAAGTGGGGCAAG	3243
DB	3185	CCCCGGCTCACAGTGGGAAACAGGGCTTCGACAGCAGCAAGTGGGGCAAG	3244
QY	3244	TCCCAGGATTTCAACCTGAGCCCTGCCACCCCTGCTGAAAAACAATCCGCA	3303
DB	3245	TCCCAGGATTTCAACCTGAGCCCTGCCACCCCTGCTGAAAAACAATCCGCA	3304
QY	3304	GAGACAGAGGAGATGGCAGGAGTTACCTGGGGGAAACAAAAACAGGATCT	3363
DB	3305	GAGACAGAGGAGATGGCAGGAGTTACCTGGGGGAAACAAAAACAGGATCT	3364
QY	3364	CCTGCTCCAGTCAGATTGGCTGACCCGCTTGGATCAGTCACCATTTGTTGG	3423
DB	3365	CCTGCTCCAGTCAGATTGGCTGACCCGCTTGGATCAGTCACCATTTGTTGG	3424
QY	3424	GGAGAGCAGCTTCACAGCTGGGTGAGAGGGGTGGCGAGCCCTTCGGCCCTC	3483
DB	3425	GGAGAGCAGCTTCACAGCTGGGTGAGAGGGGTGGCGAGCCCTTCGGCCCTC	3484
QY	3484	CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGCCCAAACCTCAGGTTCA	3543
DB	3485	CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGCCCAAACCTCAGGTTCA	3544
QY	3544	GGTTCAGCAGGATATGCCCCGCGTAGGTTAAGGGGGCCCTCTAAACCCCT	3603
DB	3545	GGTTCAGCAGGATATGCCCCGCGTAGGTTAAGGGGGCCCTCTAAACCCCT	3604
QY	3604	ACCTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAAAGATGCTCAGCC	3663
DB	3605	ACCTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAAAGATGCTCAGCC	3664
QY	3664	CCCTGGTAGAATACACACACTTTTTCAGGTTTGTGAAACACAGGTCCTGA	3723
DB	3665	CCCTGGTAGAATACACACACTTTTTCAGGTTTGTGAAACACAGGTCCTGA	3724
QY	3724	GGTTTCAGCCAGGACCAAGAGAGTGTGAAGTGAAGTGTCTCAGTCCCAGAC	3783
DB	3725	GGTTTCAGCCAGGACCAAGAGAGTGTGAAGTGAAGTGTCTCAGTCCCAGAC	3784
QY	3784	GCCCTTTTGTGCTGGCTACACACTCTTCCCGCAGACAGAGGCCCCCGAG	3843
DB	3785	GCCCTTTTGTGCTGGCTACACACTCTTCCCGCAGACAGAGGCCCCCGAG	3844
QY	3844	CCAGCATGCCCCAGATCTGCTGGCATCAGTATCCCTCATCTGTATAAGGTGA	3903
DB	3845	CCAGCATGCCCCAGATCTGCTGGCATCAGTATCCCTCATCTGTATAAGGTGA	3904
QY	3904	GCAGGATATGCTGCACAGGAACAGTCTGTGGATGGAATCATCATGCTTAAG	3963
DB	3905	GCAGGATATGCTGCACAGGAACAGTCTGTGGATGGAATCATCATGCTTAAG	3964
QY	3964	CAGAGAGAGACGTCCGGGGCCCCAGGCCCACTATCAGTGTCCAGCGTGTGGT	4023

[illegible]

```

RESULT 4
US-10-283-975A-294
; Sequence 294, Application US/10283975A
; Publication No. US20040110792A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
; FILE REFERENCE: CDS 293 PCT
; CURRENT APPLICATION NUMBER: US/10/283,975A
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/340,938
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/338,997
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/340,081
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/341,012
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 294
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: HUMAN
US-10-283-975A-294

```

```

Query Match      99.6%; Score 4578.4; DB 19; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

```

Query Match      99.6%; Score 4578.4; DB 19; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 5 GGGGACTGTGCGGTGTGGAAAGTGTAGCTGTGA-AGGTGACTCTGTACCAATTGAGG 63
Db 5 GGGGACTGTGCGGTGTGGAAAGTGTAGCTGTGAAGAGTGTGACTCTGTACCAATTGAGG 64
QY 64 ATGTTTGAGGATGAGTATGTGTGGCAGAGGCACACATAAAGCAGCAGAGACCTTTGGCC 123
Db 65 ATGTTTGAGGATGAGTATGTGTGGCAGAGGCACACATAAAGCAGCAGAGACCTTTGGCC 124
QY 124 CCTGCTTTTCCCCCAACCAAGGCTGACCTGTGTCTCCAGGCTTGGGATTTCTAAGT 183
Db 125 CCTGCTTTTCCCCCAACCAAGGCTGACCTGTGTCTCCAGGCTTGGGATTTCTAAGT 184
QY 184 GACCTGCTCTGTGTGTGTGTCTCTCAGGATGAGCAGACAGCCTGGGAGTGGAGTATG 243
Db 185 GACCTGCTCTGTGTGTGTGTCTCTCAGGATGAGCAGACAGCCTGGGAGTGGAGTATG 244
QY 244 GAAATGGCTGCCAGAGTGGCTTGGCTCAGAGTGGGAGAGGAACTCCCAAG 303
Db 245 GAAATGGCTGCCAGAGTGGCTTGGCTCAGAGTGGGAGAGGAACTCCCAAG 304
QY 304 CCAAGGAGAGACGCGCCCACTGGGAGAAACAGAGCTCCGTCTACAAGCTTTGAGGCC 363
Db 305 CCAAGGAGAGACGCGCCCACTGGGAGAAACAGAGCTCCGTCTACAAGCTTTGAGGCC 364
QY 364 GTGAGAGAGCCCTGTGTCTGTGGGAAAGTGGGAGATCTTGAATGACGTGATTTACCAAG 423
Db 365 GTGAGAGAGCCCTGTGTCTGTGGGAAAGTGGGAGATCTTGAATGACGTGATTTACCAAG 424
QY 424 GGCACGCGAGGAGGCTCCGAGCAGGCGCCAGCTGCCATCTTATCATCGCCAGGCT 483
Db 425 GGCACGCGAGGAGGCTCCGAGCAGGCGCCAGCTGCCATCTTATCATCGCCAGGCT 484
QY 484 GAGTGTGAGATAGCCAAAGAGTTCAGGCCCACTTTTCAGAACGCAATTTTCATCGCTGGG 543
Db 485 GAGTGTGAGATAGCCAAAGAGTTCAGGCCCACTTTTCAGAACGCAATTTTCATCGCTGGG 544
QY 544 TCCAAAAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCACAAATGTGGCCCATGCT 603
Db 545 TCCAAAAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCACAAATGTGGCCCATGCT 604
QY 604 ACAGAGGCAAAATGGCCCGTGTGTGTGGAGGAAAGCTCCAGCAAGCCGGAAG 663
Db 605 ACAGAGGCAAAATGGCCCGTGTGTGTGGAGGAAAGCTCCAGCAAGCCGGAAG 664
QY 664 AAACGGAAGAAAGAGAGTCAAAAGTCCCTGCTCATGCAAGAGTGGCTTTGGCCAAACCC 723
Db 665 AAACGGAAGAAAGAGAGTCAAAAGTCCCTGCTCATGCAAGAGTGGCTTTGGCCAAACCC 724
QY 724 CTCCCCAGGACCCCTGAGCAGGAGTGCACCAATCCAGTGCAGGAGGATGATCTCCA 783
Db 725 CTCCCCAGGACCCCTGAGCAGGAGTGCACCAATCCAGTGCAGGAGGATGATCTCCA 784
QY 784 CTCGGGCCCCATATGTTAGAAACACCCCGAGTTTACCAAGCTCTGAAGGAACACAGGC 843
Db 785 CTCGGGCCCCATATGTTAGAAACACCCCGAGTTTACCAAGCTCTGAAGGAACACAGGC 844
QY 844 CTTGGGCAACTCTGTTTAAAGCAGCTTGGGAGGCGCTACGGCCGGCTCTGCCTCGATCA 903
Db 845 CTTGGGCAACTCTGTTTAAAGCAGCTTGGGAGGCGCTACGGCCGGCTCTGCCTCGATCA 904
QY 904 GAACTCCACAACTGATCAGCCCTTGCATATGTTGAACACACAGTGTGGAAACTGCACAC 963
Db 905 GAACTCCACAACTGATCAGCCCTTGCATATGTTGAACACACAGTGTGGAAACTGCACAC 964
QY 964 CCCAGAGGAGGCGCCCTGCGCCACGCAACCCCTTCCCTATAGCAGACTGCCT 1023
Db 965 CCCAGAGGAGGCGCCCTGCGCCACGCAACCCCTTCCCTATAGCAGACTGCCT 1024
QY 1024 CATCCCTTCCCATTCACCTCTCCAGCCCTGGAACTTCACTCCTCTGGAGTCTTCTG 1083
Db 1025 CATCCCTTCCCATTCACCTCTCCAGCCCTGGAACTTCACTCCTCTGGAGTCTTCTG 1084

QY 1084 GGCAAACTGGCCTGTGTAGACAGCCAGAAACCTTGTGCTGACCCACACCTTGAGCAACTG 1143
Db 1085 GGCAAACTGGCCTGTGTAGACAGCCAGAAACCTTGTGCTGACCCACACCTTGAGCAACTG 1144
QY 1144 GCCTGTGTAGACAGTCAAAAGCCCTCCTGGGCCACACCTGGAGCCAGCTGCTGTCT 1203
Db 1145 GCCTGTGTAGACAGTCAAAAGCCCTCCTGGGCCACACCTGGAGCCAGCTGCTGTCT 1204
QY 1204 CGTGGTCCCATGAGAAAGTCTTGTGGAGAAATACCTAGTGCATGTCTTGCAAGGCAGC 1263
Db 1205 CGTGGTCCCATGAGAAAGTCTTGTGGAGAAATACCTAGTGCATGTCTTGCAAGGCAGC 1264
QY 1264 GTGAGCTCAAGCAGGCGCCACAGCTCAGCAGCTGGCCCAAGACCTGGGAGCAGCGGCC 1323
Db 1265 GTGAGCTCAAGCAGGCGCCACAGCTCAGCAGCTGGCCCAAGACCTGGGAGCAGCGGCC 1324
QY 1324 TCCAGATCCCGGAGCGCCAGCCCAAACTGAGGACAACTGAGGGTGTCTGCTCACTGAG 1383
Db 1325 TCCAGATCCCGGAGCGCCAGCCCAAACTGAGGACAACTGAGGGTGTCTGCTCACTGAG 1384
QY 1384 AAACCTAAGCCAGTGGATTATGATACCGAGAAAGTCCACTGGGCCACGCAAGCTC 1443
Db 1385 AAACCTAAGCCAGTGGATTATGATACCGAGAAAGTCCACTGGGCCACGCAAGCTC 1444
QY 1444 CGCTGGGAGAGGCTCCTTCGGAGAGGTGCAAGGATGAGGAGCAAGAGACTGGCTTC 1503
Db 1445 CGCTGGGAGAGGCTCCTTCGGAGAGGTGCAAGGATGAGGAGCAAGAGACTGGCTTC 1504
QY 1504 CAGTGCCTGTCAAAAAGGTGCGCTGGAAGTATTTTCGGGAGAGGAGTGTGCGCATGT 1563
Db 1505 CAGTGCCTGTCAAAAAGGTGCGCTGGAAGTATTTTCGGGAGAGGAGTGTGCGCATGT 1564
QY 1564 GCAGGATTGACCTCACCCAGAAATTTGCTTGTATGAGAGCTGTGAGAAAGGCGCTTGG 1623
Db 1565 GCAGGATTGACCTCACCCAGAAATTTGCTTGTATGAGAGCTGTGAGAAAGGCGCTTGG 1624
QY 1624 GTCAACATCTTTCAGAGCTGTGGAAGGTGGCTCCCTGGGCCAGTGTGCTCAAGAGCAG 1683
Db 1625 GTCAACATCTTTCAGAGCTGTGGAAGGTGGCTCCCTGGGCCAGTGTGCTCAAGAGCAG 1684
QY 1684 GGCCTGCTCCAGAGGACCGGGCCCTGTACTACTGCGGCCAGGCCCTGGAGGCTCGAA 1743
Db 1685 GGCCTGCTCCAGAGGACCGGGCCCTGTACTACTGCGGCCAGGCCCTGGAGGCTCGAA 1744
QY 1744 TACCTCCACTCAGAAAGATTTGCTGATGGGAGCTCAAAAGCTGCAACAGTCTCTGTCC 1803
Db 1745 TACCTCCACTCAGAAAGATTTGCTGATGGGAGCTCAAAAGCTGCAACAGTCTCTGTCC 1804
QY 1804 AGCGATGGGAGCCACGAGCCCTCTGTGACTTTTGGCCATGCTGTGTCTTCAACCTGAT 1863
Db 1805 AGCGATGGGAGCCACGAGCCCTCTGTGACTTTTGGCCATGCTGTGTCTTCAACCTGAT 1864
QY 1864 GGCCTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGCAAGAGCCCACTGGCT 1923
Db 1865 GGCCTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGCAAGAGCCCACTGGCT 1924
QY 1924 CCGAGGTGTGTGTGGCAGGAGCTGCGAGCCCAAGGTGATGTCTGGAGCAGCTGTGT 1983
Db 1925 CCGAGGTGTGTGTGGCAGGAGCTGCGAGCCCAAGGTGATGTCTGGAGCAGCTGTGT 1984
QY 1984 ATGATGTGACATGTCTCAACGGCTGCACCCCTTGGACTCAGTTCCTTCGAGGCGCGCTC 2043
Db 1985 ATGATGTGACATGTCTCAACGGCTGCACCCCTTGGACTCAGTTCCTTCGAGGCGCGCTC 2044
QY 2044 TGCCTCAAGATTGCGAGGAGCTTCGCGCTGTGAGGAGATCCCAACCTCTCTGCGCCCT 2103
Db 2045 TGCCTCAAGATTGCGAGGAGCTTCGCGCTGTGAGGAGATCCCAACCTCTCTGCGCCCT 2104
QY 2104 CTCAAGCCAGGCGCATCCCAAGAGGCGCTGAGAAAGAGCCCATCCACCGGCTGTCTGCA 2163
Db 2105 CTCAAGCCAGGCGCATCCCAAGAGGCGCTGAGAAAGAGCCCATCCACCGGCTGTCTGCA 2164
QY 2164 GCGAGCTGGGAGGAAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT 2223

1502 TCCAGTGCCTGTCAAAAAGGTGCGCTGGAAGTATTTTCGGCGAGAGGAGCTGATGGCAT 1561
1553 TCCAGTGCCTGTCAAAAAGGTGCGCTGGAAGTATTTTCGGCGAGAGGAGCTGATGGCAT 1612
1562 GTGCGAGATTGACCTTCAACCCAGAAATGTCCCTTTGTATGGAGCTGTGAGAGAAGGCCCTT 1621
1613 GTGCGAGGATTGACCTTCAACCCAGAAATGTCCCTTTGTATGGAGCTGTGAGAGAAGGCCCTT 1672
1622 GGGTCAACATCTTCAATGGAGCTGTGGAAGGTGGCTCCCTGGCGCAGAGCTGGTCAAGGAGC 1681
1673 GGGTCAACATCTTCAATGGAGCTGTGGAAGGTGGCTCCCTGGCGCAGAGCTGGTCAAGGAGC 1732
1682 AGGGCTGTCTCCAGAGGACCGGGGCCCTGTACTACTCTGGCGCAGGCGCTTGAGGGGTCTGG 1741
1733 AGGGCTGTCTCCAGAGGACCGGGGCCCTGTACTACTCTGGCGCAGGCGCTTGAGGGGTCTGG 1792
1742 AATACCTCACTCAAGAGGATTTCTGCATGGGGAAGTCAAGAGCTGACAAAGTGTCTCTGT 1801
1793 AATACCTCACTCAAGAGGATTTCTGCATGGGGAAGTCAAGAGCTGACAAAGTGTCTCTGT 1852
1802 CCAGCGATGGGAGCCACGCGACCTCTGTGACTTTTGGCCATGCTGTGTCTTCAACCTG 1861
1853 CAGCGATGGGAGCCACGCGACCTCTGTGACTTTTGGCCATGCTGTGTCTTCAACCTG 1912
1862 ATGGCTGGGAAAGTCTTTGCTCACAGGGGACTACATCCCTGGCGCAGAGACCCACATGG 1921
1913 ATGGCTGGGAAAGTCTTTGCTCACAGGGGACTACATCCCTGGCGCAGAGACCCACATGG 1972
1922 CTCGGAGGTGTGTGGCAGAGCTGGGAGCCGAGCCAGGTGATGTCTGGAGCAGCTGCT 1981
1973 CTCGGAGGTGTGTGGCAGAGCTGGGAGCCGAGCCAGGTGATGTCTGGAGCAGCTGCT 2032
1982 GTATGATGTGCACATGTCTCAACGGCTGCCACCTGTGACTCAGTTCTTCCGAGGGCCGC 2041
2033 GTATGATGTGCACATGTCTCAACGGCTGCCACCTGTGACTCAGTTCTTCCGAGGGCCGC 2092
2042 TCTGCTCAAGATTGCGAGCAGCTTCGCTGTGAGGAGATCCACCTCTCTCGGCC 2101
2093 TCTGCTCAAGATTGCGAGCAGCTTCGCTGTGAGGAGATCCACCTCTCTCGGCC 2152
2102 CTCTCACAGCCAGGCCATCCAGAGGGCTGAGGAAGAGCCCATCCAGCGGTGTCTG 2161
2153 CTCTCACAGCCAGGCCATCCAGAGGGCTGAGGAAGAGCCCATCCACCGGTGTCTG 2212
2162 CAGCGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGAGCC 2221
2213 CAGCGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGAGCC 2272
2222 CTTGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACCACC 2281
2273 CTTGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACCACC 2332
2282 AGACCTCTCATGCCAGCGAGAGCTTTCCGCAAGGCCCCAGGGCCCGCGCCAGCTG 2341
2333 AGACCTCTCATGCCAGCGAGAGCTTTCCGCAAGGCCCCAGGGCCCGCGCCAGCTG 2392
2342 AGGAGACAACAGCGAGAGCCCTTAAGTCCAGCTCCTCTCCACACAGAGCCCCAGAGC 2401
2393 AGGAGACAACAGCGAGAGCCCTTAAGTCCAGCTCCTCTCCACACAGAGCCCCAGAGC 2452
2402 CAAAACAAGTCTCCTCCCTTGACTTTGAGCAAGGAGGTCTGGGATGTGGGAACCCCTTAC 2461
2453 CAAAACAAGTCTCCTCCCTTGACTTTGAGCAAGGAGGTCTGGGATGTGGGAACCCCTTAC 2512
2462 CTCTGCTCCTCTGGAGCGCCCTCTGCAGAAACCCGAGCTCACAGAGCGGGAAGCA 2521
2513 CTCTGCTCCTCTGGAGCGCCCTCTGCAGAAACCCGAGCTCACAGAGCGGGAAGCA 2572
2522 CCGTCCCGGAGCAGGAACTGCAGCAGCTGGAATAGAATTATCTCTCAACAGCCTGTCCC 2581
2573 CCGTCCCGGAGCAGGAACTGCAGCAGCTGGAATAGAATTATCTCTCAACAGCCTGTCCC 2632
2582 AGCCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATTCGACAGCCTCT 2641

2633 AGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTCAGCATTCGACAGCCTCT 2692
2642 CCCTGTCCGATGACAGTTCAGAGAACCCCATCAAAAGCCCTCTCAAAAGCTCGCGGACACC 2701
2693 CCCTGTCCGATGACAGTTCAGAGAACCCCATCAAAAGCCCTCTCAAAAGCTCGCGGACACC 2752
2702 TGAGCTCAGGCGTACACTCTCTGGAGCAGCAGGCGCAGGCTCGAAGCTCAGCTGGAACA 2761
2753 TGAGCTCAGGCGTACACTCTCTGGAGCAGCAGGCGCAGGCTCGAAGCTCAGCTGGAACA 2812
2762 TGTGCTCGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAGTCC 2821
2813 TGTGCTCGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAGTCC 2872
2822 AATAACAGTCTCTTAATGGTGAACACCTTGACATCCGGGAGTTCACCGGCTCAAGTGG 2881
2873 AATAACAGTCTCTTAATGGTGAACACCTTGACATCCGGGAGTTCACCGGCTCAAGTGG 2932
2882 GAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTTCAGCTTGGTCAACA 2941
2933 GAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTTCAGCTTGGTCAACA 2992
2942 AAGACGGCAGCCTGTTTCGCTACGATGGAGGTGCCAGACTCGGGCATTCGACCTGCAAGT 3001
2993 AAGACGGCAGCCTGTTTCGCTACGATGGAGGTGCCAGACTCGGGCATTCGACCTGCAAGT 3052
3002 GCACACTCGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTGAAGCATGGCCAGCTGG 3061
3053 GCACACTCGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTGAAGCATGGCCAGCTGG 3112
3062 AGAACAGGCCCTTAACCTCGCTCCACCGCGGCTCCACACTGCGGGAAGCAGCTTCC 3121
3113 AGAACAGGCCCTTAACCTCGCTCCACCGCGGCTCCACACTGCGGGAAGCAGCTTCC 3171
3122 TGCTCGTGCAGATGTGCTCCCTGAAAAACAGAGCTCAGCGCTTCCGAGGGAATTGCCAG 3181
3172 TGCTCGTGCAGATGTGCTCCCTG-AAACACAGGCTCAGCGCTTCCGAGGGAATTGCCAG 3230
3182 CCCCCCGCTCACAGTGGGAAACAGGCGCTCGCAGCAGCAGGTGGGGCAGCAGAAATG 3241
3231 -CCCCCGCTCAGCAGTGGGAAACAGGCGCTCGCAGCAGCAGGTGGGGCAGCAGAAATG 3289
3242 CTTCCAGGATTTTCACACTGAGCCTGCCCCACCTCTGCTGAAAAAACA-TCCGCCACGT 3300
3290 CTTCCAGGATTTTCACACTGAGCCTTGCCCCACCTCTGCTGAAAAAACA-TCCGCCACGT 3349
3301 GAAAGACAGAGGAGGATGGCAGAGTTTACTGGGGGAAAAAACAACAGGGAATCTTTTCT 3360
3350 GAAAGACAGAGGAGGATGGCAGAGTTTACTGGGGGAAAAAACAACAGGGAATCTTTTCT 3404
3361 GCCCTGCTCCAGTTCGAGTTCGCTGACCGCTTGGATCAGTGACCATTTGTTGGCAGAC 3420
3405 TGCCCTGCTCCAGTTCGAGTTCGCTGACCGCTTGGATCAGTGACCATTTGTTGGCAGAC 3464
3421 AGGGAGAGCAGCTTTCAGCCTGGGTGAGAGGGGTGGCGCAGCCTTCGGCCCTTCACC 3480
3465 AGGGAGAGCAGCTTTCAGCCTGGGTGAGAGGGGTGGCGCAGCCTTCGGCCCTTCACC 3524
3481 CTCAGGCTGCTGTGAGAGTTCAGTGTGTAAGGGCCCAAACTCAGGTTTCAGTGCAGAA 3540
3525 CTCAGGCTGCTGTGAGAGTTCAGTGTGTAAGGGCCCAAACTCAGGTTTCAGTGCAGAA 3584
3541 CCAGTTCAGCAGGATGCTCCCGCTAGTGTAAAGGGCCCTCTCTAAACCCCTTGTCTGCG 3600
3585 CCAGTTCAGCAGGATGCTCCCGCTAGTGTAAAGGGCCCTCTCTAAACCCCTTGTCTGCG 3644
3601 CTCACCTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAAAGAAATGCCCTGACCTGGGAAG 3660
3645 CTCACCTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAAAGAAATGCCCTGACCTGGGAAG 3704
3661 GCTCCCTGGTGAATAACCAACATTTTTCAGTGTGTTGCAACACAGGCTCTCAGTTGACC 3720

Db 781 AGCAGATGCTCATCCCTTCCATTCCACCCCTCCAGCCCTGGAACCTCACCCCTCTG 840
Qy 1072 GAGTCTTCTCTGGCAAACTGSCCTGTGTAGACAGCCAGAAACCTTGGCTGACCCACAC 1131
Db 841 GAGTCTTCTCTGGCAAACTGSCCTGTGTAGACAGCCAGAAACCTTGGCTGACCCACAC 900
Qy 1132 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTCGCTGGGCCACACCTCGAGGCC 1191
Db 901 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTCGCTGGGCCACACCTCGAGGCC 960
Qy 1192 AGCTGCCCTGTCTGTGGTCCCATAGAAATTTTCTGTGGAGAAATCACTAGTGCATGCT 1251
Db 961 AGCTGCCCTGTCTGTGGTGGCTGCCCATGAGAAATTTTCTGTGGAGAAATCACTAGTGCATGCT 1020
Qy 1252 CTGCAAGGCAGGCTGAGCTCAAGCCAGGCCCAAGCCTTGAACAGCTGGCCAAAGCCTGG 1311
Db 1021 CTGCAAGGCAGGCTGAGCTCAAGCCAGGCCCAAGCCTTGAACAGCTGGCCAAAGCCTGG 1080
Qy 1312 CGAGCACGGGGCTCCAGATCCCGGAGGCCAGCCCCAAATCTGAGGACAAACGAGGGTGTCT 1371
Db 1081 CGAGCACGGGGCTCTAGATCCCGGAGGCCAGCCCCAAATCTGAGGACAAACGAGGGTGTCT 1140
Qy 1372 CTGCTCACTGAGAAACTCAAGCCAGTGGATTATGATACCGAGAAAGATCCACTGGGCC 1431
Db 1141 CTGCTCACTGAGAAACTCAAGCCAGTGGATTATGATACCGAGAAAGATCCACTGGGCC 1200
Qy 1432 AGGCACCAAGCTCCGCTGGGAGAGAGCTCTTCCGAGAGGTTGACAGGATGGAGACAAG 1491
Db 1201 AGGCACCAAGCTCCGCTGGGAGAGAGCTCTTCCGAGAGGTTGACAGGATGGAGACAAG 1260
Qy 1492 CAGACTGGCTTCCAGTGGCTGTCAAAAAGGTGGCCCTGGAAGTATTTGGGACAGGAG 1551
Db 1261 CAGACTGGCTTCCAGTGGCTGTCAAAAAGGTGGCCCTGGAAGTATTTGGGACAGGAG 1320
Qy 1552 CTGATGGCATGTGCAAGATTTGACCTCACCAGAAATTTGCCCTTTGTATGGAGCTGTGAGA 1611
Db 1321 CTGATGGCATGTGCAAGATTTGACCTCACCAGAAATTTGCCCTTTGTATGGAGCTGTGAGA 1380
Qy 1612 GAAAGGCTTGGGTCAACATCTTCAATGAGCTGTGGAAGTGGCTCCCTGGGCCAGCTG 1671
Db 1381 GAAAGGCTTGGGTCAACATCTTCAATGAGCTGTGGAAGTGGCTCCCTGGGCCAGCTG 1440
Qy 1672 GTCAAGGAGCAGGCTGTCTCCAGAGACCGGGCCCTGTACTACTGGGCCAGGCCCTG 1731
Db 1441 GTCAAGGAGCAGGCTGTCTCCAGAGACCGGGCCCTGTACTACTGGGCCAGGCCCTG 1500
Qy 1732 GAGGCTCTGGAATACCTCACTCAGCAAGGATTTGCTATGGGGAGCTCAAAAGCTGACAAC 1791
Db 1501 GAGGCTCTGGAATACCTCACTCAGCAAGGATTTGCTATGGGGAGCTCAAAAGCTGACAAC 1560
Qy 1792 GTGCTCTGTTCAGCGATGGAGGCCACGAGCCCTCTGTGACTTTGGGCAATGCTGTGTGT 1851
Db 1561 GTGCTCTGTTCAGCGATGGAGGCCACGAGCCCTCTGTGACTTTGGGCAATGCTGTGTGT 1620
Qy 1852 CTTCAACTGTATGGCTGGGAAGTCTTTGCTCACAGGGAGTACATCCCTGGGCACAGAG 1911
Db 1621 CTTCAACTGTATGGCTGGGAAGTCTTTGCTCACAGGGAGTACATCCCTGGGCACAGAG 1680
Qy 1912 ACCCAATGGCTCCGAGAGTGGTGTGGGACAGAGCTGCGACGCCAAAGTGGATGTCTGG 1971
Db 1681 ACCCAATGGCTCCGAGAGTGGTGTGGGACAGAGCTGCGACGCCAAAGTTCGACGTCTGG 1740
Qy 1972 AGCAGCTGTGTATGATGCTGACATGCTCAACGGCTGGCCACCCCTGGAATCAGTTCTTC 2031
Db 1741 AGCAGCTGTGTATGATGCTGACATGCTCAACGGCTGGCCACCCCTGGAATCAGTTCTTC 1800
Qy 2032 CGAGGGCCGCTCTGCTCAAGATTTCCAGCGAGCTCCGCTGTGAGGGAGATCCACCC 2091
Db 1801 CGAGGGCCGCTCTGCTCAAGATTTCCAGCGAGCTCCGCTGTGAGGGAGATCCACCC 1860
Qy 2092 TCCTCGGCCCTCTCACAGCCCAAGCCATCCAAGAGGGGCTGAGGAAAGAGCCCATCCAC 2151
Db 1861 TCCTCGGCCCTCTCACAGCCCAAGCCATCCAAGAGGGGCTGAGGAAAGAGCCCATCCAC 1920

Qy 2152 CGCGTGTCTGCAGCGGAGCTCGGAGGGAAGGTGAACCGGGCACTACACAAAGTGGAGGT 2211
Db 1921 CGCGTGTCTGCAGCGGAGCTCGGAGGGAAGGTGAACCGGGCACTACACAAAGTGGAGGT 1980
Qy 2212 CTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCCGCCAAATCAAGCC 2271
Db 1981 CTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCCGCCAAATCAAGCC 2040
Qy 2272 AATTACCAACAGACCCCTTCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCC 2331
Db 2041 AATTACCAACAGACCCCTTCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCC 2100
Qy 2332 CGGCCAGCTGAGGAGAGCAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACACAGAG 2391
Db 2101 CGGCCAGCTGAGGAGAGCAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACACAGAG 2160
Qy 2392 CCCCAGAGCCAAACAGTCTCTCCCTTGAATTTTGAACAAGGAGAGTCTGGGATGTGG 2451
Db 2161 CCCCAGAGCCAAACAGTCTCTCCCTTGAATTTTGAACAAGGAGAGTCTGGGATGTGG 2220
Qy 2452 GAAACCTTACCTCTGTCTCTCGCTGAGAGCCCTTCGACAGAAACCCAGCTCACCAGAG 2511
Db 2221 GAAACCTTACCTCTGTCTCTCGCTGAGAGCCCTTCGACAGAAACCCAGCTCACCAGAG 2280
Qy 2512 CGGAAAGCAACCGTCCCGGAGCAGGAACTGACAGAGCTGGAAATAGAATTTTCTCTCAAC 2571
Db 2281 CGGAAAGCAACCGTCCCGGAGCAGGAACTGACAGAGCTGGAAATAGAATTTTCTCTCAAC 2340
Qy 2572 AGCTGTCTCCAGCCATTTTCTCTGAGAGCAGAGCAAAATTTCTCTCGCTCGACATC 2631
Db 2341 AGCTGTCTCCAGCCATTTTCTCTGAGAGCAGAGCAAAATTTCTCTCGCTCGACATC 2400
Qy 2632 GACAGCTCTCTCCCTGTGAGATGACAGTGAAGAAACCCATCAAGGGCTCTCAAAAGCTCG 2691
Db 2401 GACAGCTCTCTCCCTGTGAGATGACAGTGAAGAAACCCATCAAGGGCTCTCAAAAGCTCG 2460
Qy 2692 CGGGACACCTTGAGCTCAGGCGTACATCTCTGGAGCAGCCAGGGCTCGAGGCTCGAGCTCC 2751
Db 2461 CGGGACACCTTGAGCTCAGGCGTACATCTCTGGAGCAGCCAGGGCTCGAGGCTCGAGCTCC 2520
Qy 2752 AGCTGGAACATGTGTCTGGCCGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGT 2811
Db 2521 AGCTGGAACATGTGTCTGGCCGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGT 2580
Qy 2812 GTGAAGTCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCGGGAGTTCACCCGG 2871
Db 2581 GTGAAGTCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCGGGAGTTCACCCGG 2640
Qy 2872 GTCAAGTGGGAGACATCGCCATCAGCAGCAGATCCAGCTCGAGCTTCAGC 2931
Db 2641 GTCAAGTGGGAGACATCGCCATCAGCAGCAGATCCAGCTTCAGCAGCTTCAGC 2700
Qy 2932 TTGGTCAACAAAGACGGGACGCTTTCTCGCTACGACATGGAGGTGCCAGACTCGGGCATC 2991
Db 2701 TTGGTCAACAAAGACGGGACGCTTTCTCGCTACGACATGGAGGTGCCAGACTCGGGCATC 2760
Qy 2992 GACCTGAGTGACATCTGGCCCTGTGATGGAGCTTTGCTGAGAGCTGAGGGTCAAGCAT 3051
Db 2761 GACCTGAGTGACATCTGGCCCTGTGATGGAGCTTTGCTGAGAGCTGAGGGTCAAGCAT 2820
Qy 3052 GGCCAGCTGGAGAACAGGCCCTAA 3075
Db 2821 GGCCAGCTGGAGAACAGGCCCTAA 2844

RESULT 7

US-10-087-192-884

; Sequence 884, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.

; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR

; FILE OF INVENTION: CANCER

; FILE REFERENCE: 529452000122

; CURRENT APPLICATION NUMBER: US/10/087,192

; CURRENT FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: US 09/747,377

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: US 09/798,586

; PRIOR FILING DATE: 2001-03-02

; NUMBER OF SEQ ID NOS: 2059

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 884

; LENGTH: 2828

; TYPE: DNA

; ORGANISM: Mus musculus

US-10-087-192-884

Query Match 44.4%; Score 2042.6; DB 13; Length 2828;

Best Local Similarity 83.4%; Pred. No. 0;

Matches 2377; Conservative 0; Mismatches 444; Indels 28; Gaps 4;

Qy	232	ATGCGAGTGTGAAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAG	291
Db	1	ATGCGCGTGTGAAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAG	60
Qy	292	GAATCTCCCAAGCCAAAGGAGAGACCGCCACTGCGGGAGAAACAGAGCTCCGTCTAC	351
Db	61	GAGCTTGCCAAAGCCAAAGGAGAGACAGTCACTGCGGGAAGAGCAGAGCTGCATCTTC	120
Qy	352	AAGCTTGAGGCGGTGGAGAGAGAGCCCTGTGTTCTGCGGAAGATGGGAGATCTGAATGAC	411
Db	121	AAGCTTGAGGCGGTGGAGAGAGAGCCCTGTGTTCTGCGGAAGATGGGAGATCTTAACGAC	180
Qy	412	GTGATTACCAAGGCGCAGACCAAGAGAGGTCCGAGCAGGCGCAGCTGCCATCTCTATC	471
Db	181	GTGATTACCAAGGCGCAGACCAAGAGAGGTCTGAGGAGAGACCAAGGCGCATCTCCATC	240
Qy	472	ATCCGCCAGGCTGAGTGTGAAATAGCAAGAGTTCAGCCCCACCTTTTCAGAAACGCAAT	531
Db	241	ATCCGCCAGGCTGAATGTGAAATAGCAAGAGTTCAGCCCCACCTTTTCAGAGCGCAAT	300
Qy	532	TTGATCGTGGTCCAAACAGTACAGCCAGTCCAGAGTCTTGATCAGATCCCAACAAAT	591
Db	301	TTGATCGGCGGTCAACAGCAGTACAGCCAGTCTGAGAGTCTCGATCAAAATCCCCAACAAAT	360
Qy	592	GTGCGCCATGCTACAGAGGGCAATATGCCCGTGTGTTGGAAGGGAACGTCGCGAGC	651
Db	361	GTGCGCCATGCAACTGAAGGCAATATGCCCGTGTGTTGCGGAGGGAACGTCACGCGC	420
Qy	652	AAAGCCCGGAAGAAACGGAAGAGAGCTCAAGTCCCTGGCTCATGCGAGAGTGGCC	711
Db	421	AAAGCCCGGAAGAAACGTAAGAGAGAGGTGCAAGTCACTGGCCAGGCGAGTGGCC	480
Qy	712	TTGGCCAAACCCCTCCCGAGACCCCTAGCAGAGAGAGTGCATCATCCAGTCAGAGAG	771
Db	481	TTAGCCAAAGCCCTCCCGAGACCCCTAGCAGAGAGAGTGTACATATCCAGTACAGGAA	540
Qy	772	GATGAGTCTCACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCCCTG	831
Db	541	GATGAGTCTCACTAGGCAACCTTATGCGCAAGAAATGTCTCCAGTTTCAACAGCCCTG	600
Qy	832	AAGGAACAGAGCCCTTGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGCTTACGCCCGGCT	891
Db	601	GGGGGACAGAGCCCTTGGCAACTCTGTTTAAAGAAACAGATGAAGGCTTGGCAACGGTA	660
Qy	892	CTGCGCTCGATCAGAACTCCAAACTGATCAGCCCTTTGCAATGTCTGAACCAAGTGTGG	951
Db	661	CTGCGCTCGAACAGAACTCCAAACTGATCAGCCCTTTGCAATGTCTGAACCAAGTGTGG	720
Qy	952	AAATGCAACCAACCCAGGAGCGAGGGCCCTGCGCCCTGCCAGCAGCCCTTCCCTAT	1011
Db	721	AAATGCAACCAACCCAGGAGCGAGGGCCCTGCGCCCTGCCAGCAGCCCTTCCCTAT	780

Db 1860 CCACCTTCTGCGCACCCCTCAAGAGAGGGCTGAGGAAAGAGCCC 1919
Qy 2146 ATCCACGGGTGCTGCGAGGAGCTGGAGGAGGTGAACCGGCGACTACAGCAAGTG 2205
Db 1920 GTCCACGAGCATCTGCCATGGAGCTTCGGAGGAAAGTGGCAAGGCATACAGGAAGTG 1979
Qy 2206 GGAGGTCTGAAGAGCCCTTGGAGGGGAGATATAAAGAACCAAGACATCACCGCAAT 2265
Db 1980 GGAGGTCTGAAGAGCCCTTGGAGGAGGATATAAAGAACCAAGACATCACCGCAAGAC 2039
Qy 2266 CAAAGCCAAATACCAAGAGCCCTTCCATGCCAGCGGAGAGCTTTTCGCAAGGGCCCCA 2325
Db 2040 CAAAGCCAACTGCGACCAAGAGCCCTTACCTACTCGCGGAGAGA----- 2080
Qy 2326 GGGCCCCGCGAGCTGAGGAGACACAGGAGAGGCCCCCTAAGTCCAGGCTCTCTCCCA 2385
Db 2081 --GAACCCAGCCCAAGGCCAACACAGAGCGGGGCTCTGAGCCTCAGCCTCTCTACCG 2138
Qy 2386 CCAGAGCCCCCAGAGCAAAAGTCTCTCTGCTGACCTTGAGCAAGGAGGCTGGG 2445
Db 2139 CAGAAACCAAGAGCCCTTGGAGGAGGATATAAAGAACCAAGACATCACCGCAAGAC 2198
Qy 2446 ATGTGGAAACCTTACTCTGCTCTCTCTGAGCCAGCCCTTGCAGAAACCCAGCTCA 2505
Db 2199 ACATGGGAACCCCTGCTCTCTCTCTGAGCCAGCCCACTGCGCAAGGCCCCAGCTTC 2258
Qy 2506 CCAGAGCGGAAGCAACCTCGGAGCAGGAACTGCGAGCAGCTGGAATAGAAATATTC 2565
Db 2259 CCAGAGCGGAGGCAACCTTGCAGAGCTGAGCTACAGCAACTGGAGATAGAACTGTTT 2318
Qy 2566 CTCAACAGCCTGTCAGAGCAATTTCTGAGGAGCAGGAGCAAAATCTCTCGTCCCTC 2625
Db 2319 CTCAACAGCCTGTCAGAGCGCTCTCTCTGAGGAGCAGGAGCAAAATCTCTCTGCTC 2378
Qy 2626 AGCATCGAAGCCTCTCTGCTGAGTGAAGCAAGTGAAGAAACCCATCAAGGCTCTCAA 2685
Db 2379 AGCATCGAAGCCTCTCTGCTGAGTGAAGCAAGTGAAGAAACCCATCAAGGCTCTCAA 2438
Qy 2686 AGCTCGCGGACACCTGAGCTCAGGCTACACTCTGAGCAGCCAGGCGCAGGCTCGA 2745
Db 2439 AGCTCAGGGAACCCCTGAGTTCTGCGGTGCACTCTTGGAAACAGCAAGCTGAGGCAAGA 2498
Qy 2746 AGCTCAGCTGGAACATGCTGTGCGCGCGGCGGCCACCGACACCCCAAGCTATTTC 2805
Db 2499 ACCTGAGCTGAGCAGCGCGCTGCGCGGCGGCGGCTACTGACATCCCGAGCTACTTC 2558
Qy 2806 AATGCTGGAAGTCCAAATACAGTCTCTTAATGTTGAACACCTGCAATCCGGAGTTC 2865
Db 2559 AACGGGGTCAAGGTCCAGATCCAGTCTCTCAATGGCGAAACCTGCATATCCGGAAATTC 2618
Qy 2866 CACCGGTCAAAGTGGAGACATCGCCACTGCGCATCAGCAGCCAGATCCAGCTCGAGCC 2925
Db 2619 CACCGGTCAAGGTGGAGACATGCCCACCGGCATCAGCAGCCAGATCCAGCCACAGCT 2678
Qy 2926 TTGAGCTTGGTCAACAAAGACGGGAGCCTGTTTCTGCTACGATGAGGTGCGAGCTCG 2985
Db 2679 TTGAGCTTGGTCAACAAAGTGAACAGCCTGTTTCTGCTATGATGAGGTGCGAGCTCG 2738
Qy 2986 GGCATGACCTGAGTGCACTGCGCCCTGATGAGGCTTCGCTGAGGCTGAGGGTTC 3045
Db 2739 GGCATGACCTGAGTGCACTGCGCCCTGATGAGGCTTCGCTGAGGCTGAGGGTTC 2798
Qy 3046 AAGCATGGCAGCTGGAGAACAGGCCCTA 3074
Db 2799 AAGCATGGTCTGAGTGAGAACCGACCCCTA 2827

RESULT 8
US-10-087-192-886
; Sequence 886, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:

; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 886
; LENGTH: 73967
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-886

Query Match 33.9%; Score 1556.2; DB 13; Length 73967;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1650; Conservative 0; Mismatches 23; Indels 10; Gaps 7;

Qy 2909 AGATCCCAAGCTGCAGCCTTTCAGCTTGGTCAACAAAGACGGGAGCCTGTTGCTACGACA 2968
Db 62243 AGATCCCAAGCTGCAGCCTTTCAGCTTGGTCAACAAAGACGGGAGCCTGTTGCTACGACA 62302

Qy 2969 TGGAGTGCCAGACTCGGGCATTCGACCTGCGAGTGACACTGCGCCCTGATGGCAGCTTCG 3028
Db 62303 TGGAGTGCCAGACTCGGGCATTCGACCTGCGAGTGACACTGCGCCCTGATGGCAGCTTCG 62362

Qy 3029 CCTGGAGCTGAGGGTCAAGCATGCGCAGCTGAGGAAACAGCCCTTAACCCCTGCCCTCCAC 3088
Db 62363 CCTGGAGCTGAGGGTCAAGCATGCGCAGCTGAGGAAACAGCCCTTAACCCCTGCCCTCCAC 62422

Qy 3089 CGCCGGCTCCACACTGCGGAAAGCAGCCTTCTGCTCGTGCGAGTGCCTGCTGCCCTGAAA 3148
Db 62423 CGCCGGCTCCACACTGCGGAAAGCAGCCTTCTGCTCGTGCGAGTGCCTGCTGCCCTGAAA 62480

Qy 3149 ACACAGGCTCAGCGTTCGCCAGGGGATTCGACAGCCCGCCCGGCTCAGTGGGAACAGGG 3208
Db 62481 ACACAGGCTCAGCGTTCGCCAGGGGATTCGACAGCCCGCCCGGCTCAGTGGGAACAGGG 62539

Qy 3209 CCTGCAGCAGCAGGCTGGGGGCAAGCAGAGTGCCTCCAGAGTTCACACTGAGCCCT 3268
Db 62540 CCTGCAGCAGCAGGCTGGGGGCAAGCAGAGTGCCTCCAGAGTTCACACTGAGCCCT 62599

Qy 3269 GCGCCACCTGCTGAAAAAACA-TCCGCCACGTGAAGAGACAGAGGAGGATGGCAGGAG 3327
Db 62600 GCGCCACCTGCTGAAAAAACA-TCCGCCACGTGAAGAGACAG-AGGAGGATGGCAGGAG 62658

Qy 3328 TTACTTGGGAAAAACAAACAGGGATCTTTTCTGCGCCCTGCTCCAGTGCAGTGGCCCTGA 3387
Db 62659 TTACTTGGGAAAAACAAACAGG-ATCTTCTGCTGCTGCTCCAGTGCAGTGGCCCTGA 62714

Qy 3388 CCGCTTGGATCAGTGACCATTTGTCAGACAGGGGAGAGCAGCTTCCAGCCTGGGTC 3447
Db 62715 CCGCTTGGATCAGTGACCATTTGTCAGACAGGGGAGAGCAGCTTCCAGCCTGGGTC 62774

Qy 3448 AGAAGGGTGGGCGAGCCCTTTCGCGCCCTCACCCTCCAGGCTGCTGTGAGAGTGTCAAGT 3507
Db 62775 AGAAGGGTGGGCGAGCCCTTTCGCGCCCTCACCCTCCAGGCTGCTGTGAGAGTGTCAAGT 62834

Qy 3508 GTGTAAAGGGCCCAAACTCAGGTTTTCAGTGAGAACCAAGTTCAGAGGATGTCGCCGCCGTA 3567
Db 62835 GTGTAAAGGGCCCAAACTCAGGTTTTCAGTGAGAACCAAGTTCAGAGGATGTCGCCGCCGTA 62894

Qy 3568 GGTAAAGGGGCGCCCTCTAAACCCCTTGCCTGCGCTCACCTGCGCAGCTCAACCCCTTTGG 3627
Db 62895 GGTAAAGGGGCGCCCTCTAAACCCCTTGCCTGCGCTCACCTGCGCAGCTCAACCCCTTTGG 62954

Qy 3628 GTGTAGGGAAAAAAGATGCTGACCTTGGGAGGCTTCCCTGGTAGAATACACCACTTT 3687

Db 781 TATCAGTGTCCAGCGTGTGGTTCCCGAGAGCACAGCTCAGCATCACACTGACACTCACC 840
QY 4055 CTGCCCTGCCCTGGCGCAGAGGTACTGCGAGCGGCACCTTTTGACTCTGATGACCTCAA 4114
Db 841 CTGCCCTGCCCTGGCGCAGAGGTACTGCGAGCGGCACCTTTTGACTCTGATGACCTCAA 900
QY 4115 GCACCTTTCATGGCTGCCCTCTGGCAGGGCAGGGCAGTGACACACTGTAGGAGCATAG 4174
Db 901 GCACCTTTCATGGCTGCCCTCTGGCAGGGCAGGGCAGTGACACACTGTAGGAGCATAG 960
QY 4175 CAAAGCAGAGATGGGGTGAAGGACACAGTCTTGTAGCTGTGCCACATGATGTACTCCT 4234
Db 961 CAAAGCAGAGATGGGGTGAAGGACACAGTCTTGTAGCTGTGCCACATGATGTACTCCT 1020
QY 4235 CAAACCTCTTCCAGATTTCTCTAAGAAATAGCACCCCTTCCCATTTGCCCGAGCTAGCC 4294
Db 1021 CAAACCTCTTCCAGATTTCTCTAAGAAATAGCACCCCTTCCCATTTGCCCGAGCTAGCC 1080
QY 4295 TCTTCTCCAGGGAGTACTCAGGACTCACGTAGCATTAATCAATCAGCTGTGAATCGTCAG 4354
Db 1081 TCTTCTCCAGGGAGTACTCAGGACTCACGTAGCATTAATCAATCAGCTGTGAATCGTCAG 1140
QY 4355 GGGGTGTCTGTAGCTCTAACTCTGGGGCAGGGGACCCGAGACTCCGTGGGAGAAC 4414
Db 1141 GGGGTGTCTGTAGCTCTAACTCTGGGGCAGGGGACCCGAGACTCCGTGGGAGAAC 1200
QY 4415 TCATTCACCATCTTGCACAGACAGCTTTGTGCCAGCTGTCCACATTTGAGTCAGACTGCT 4474
Db 1201 TCATTCACCATCTTGCACAGACAGCTTTGTGCCAGCTGTCCACATTTGAGTCAGACTGCT 1260
QY 4475 CCCGGGAGAGACCCCGGCCCGCCAGCACATAAGAACTGACGCTTGGTACTGCGAGAGT 4534
Db 1261 CCCGGGAGAGACCCCGGCCCGCCAGCACATAAGAACTGACGCTTGGTACTGCGAGAGT 1320
QY 4535 CTGGGTGTAGAGAACTCTTTGTAGCAATAAAGTTTGGGGTGTAGCAAAATGTTAAAAA 4594
Db 1321 CTGGGTGTAGAGAACTCTTTGTAGCAATAAAGTTTGGGGTGTAGCAAAATGTTAAAAA 1380
QY 4595 AA 4596
Db 1381 AA 1382

RESULT 10
US-10-087-192-883
; Sequence 883, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 883
; LENGTH: 42566
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(42566)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-883

Query Match 10.4%; Score 480.2; DB 13; Length 42566;
Beat Local Similarity 74.3%; Pred. No. 1.6e-132;

Matches 687; Conservative 0; Mismatches 153; Indels 85; Gaps 3;
QY 554 ACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAATGTGGCCCATGCTACAGAGGCA 613
Db 13035 ACAGCCAGTCTGAGAGTCTCGATCAAAATCCCAACAATGTGGCCCATGCTACAGAGGCA 13094
QY 614 AAATGGCCCGTGTGTGTGGAGGGAAGCGTCGCAGCAAAAGCCGGAAGAAACCGAAGA 673
Db 13095 AAATGGCCCGTGTGTGTGGAGGGAAGCGTCGCAGCAAAAGCCGGAAGAAACCGTAGGA 13154
QY 674 AGAAGAGTCAAAAGTCCCTGGCTCATGAGAGTGGCTTGGCCAAACCCCTCCACAGA 733
Db 13155 AGAAGAGTCAAAAGTCACTGGCCAGGAGGAGTGGCTTGGCCAAAGCCCTGCCAGAA 13214
QY 734 CCCCTGAGCAGGAGAGCTGCACCATCCAG----- 763
Db 13215 CCCCTGAGCAAGAGAGCTGTACCATCCAGTACAGGTAGATCTCTGCCCGTTGCCCTGGT 13274
QY 764 ----- 774
Db 13275 CCCAGTTCATGGCGGTTTGGACTGGAGCTTGTAAATGCTGTATTTTCTGCAGGAAGAT 13334
QY 775 GAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCTCTGAAG 834
Db 13335 GAGTCTCCACTAGGCAACCTCTATGCCAGAAATGTCTCCAGTTTACCAAGCTCTGGGG 13394
QY 835 GAAACAGGCTTGGGCAACTCTGTTTTAAGCAGCTTGGGGAGGCTTACGCCGCTCTG 894
Db 13395 GAAACAGGCTTGGGCACTGTGCTTTAAGAAACAAAGATGAAGGCTTCGACCGGTACTG 13454
QY 895 CCTCGATCAGAACTCCACAAACTGATCAGCCCTTGCATGTCTGAAACACAGTGTGAAA 954
Db 13455 CCTCGACAGAACTCCACAAACTGATCAGCCCTTGCATGTCTTAAACACAGTGTGAAA 13514
QY 955 CTGCACACCCCGCAGGAGCGGCCCTGCCCCCTGCCCCAGCACCCCTTCCCCCTATAGC 1014
Db 13515 CTGCACACCCCGCAGGCGCCAGGCGGCCCGGCCAGCTCAGCCCTTCCCCCTACAGC 13574
QY 1015 AGACTGCTCATCCCTTCCGATTCACCCCTCTCCAGCCCTTGAACCTTCACCCCTCTGGAG 1074
Db 13575 GAAATGCCCCCATCTCTTCCCATTTACCCCTTGGAGCCCTTGAACCTTATATGCTGAC 13634
QY 1075 T---CTTCTGGGCAAACTGCGCTGTGTAGACAGCCAGAAACCTTCCCTGACCCA--- 1128
Db 13635 TCTGCGCTCTTGACAAACTAGCCGCTGTACGCGGCGAGCGGCTCTGCTTGGCCACCG 13694
QY 1129 CACTGAGCAAACTGGCCTGTGTAGACAGTCCAAAGCCCTTCCCTGGGCCACACCTGGAG 1188
Db 13695 CATCTAAGCCCAACTGGCCCATGGAGACAGTCAAGAGCGCTGCTGGGCCACACCTGGAG 13754
QY 1189 CCCAGCTGCTGTCTCGTGGTGCCCATGAGAAAGTTTCTGTGGAGGAATACCTAGTGCAT 1248
Db 13755 TCCAGCTGCCCTCTCGGGGTGCCCTAGAAAAGGTTCCCGTGGAGGAATACCTGCTGCAT 13814
QY 1249 GCTCTGAAAGCAGCGTGTAGCTCAAGCCAGGCGCCACACAGCTTGCAGCGCTGGCCAAAGACC 1308
Db 13815 GCGCTCCAAAGAGTGTGAGCTCAGGCCAGGCCACACAGCTTGGCCAGCTGGCTAAGACA 13874
QY 1309 TGGGAGCAGCGGGCTCCAGATCCCGGAGCGCCAGCCCAAACTGAGGACAAAGAGGT 1368
Db 13875 TGGTCTCGGGAAGCGCCAGCTGCAGAGGCTCGGCCCGGAACTGAGGACAAAGAGGGG 13934
QY 1369 GTCTGTCTCACTGAGAACTCAAGC 1393
Db 13935 GTCTGTCTTACTGAGTAACTAACC 13959

RESULT 11
US-10-108-260A-1754/c
; Sequence 1754, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE


```
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1754
; LENGTH: 2475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1754

Query Match          7.6%; Score 349.4; DB 17; Length 2475;
Best Local Similarity 99.7%; Pred. No. 1.1e-93;
Matches 350; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2200 CAAATGAGGAGTCTGAAGAGCCCTTGGAGGGGAGAAATAAAGAACCAAGACATCCACCG 2259
Db 2179 CTAGTGGAGGTCTGAAGAGCCCTTGGAGGGGAGAAATAAAGAACCAAGACATCCACCG 2120

Qy 2260 CCAATCAAGCAATTAACACAGACCCTTCATGCCAGCCGAGAGAGCTTTCGCCAAGG 2319
Db 2119 CCAATCAAGCAATTAACACAGACCCTTCATGCCAGCCGAGAGAGCTTTCGCCAAGG 2060

Qy 2320 GCCCCAGGGCCCCGCCAGCTGAGGAGACAACAGCAGAGCCCTTAAGCTCCAGCCTCT 2379
Db 2059 GCCCCAGGGCCCCGCCAGCTGAGGAGACAACAGCAGAGCCCTTAAGCTCCAGCCTCT 2000

Qy 2380 CTCACCACAGAGCCCCCAGAGCCAAACAAGTCTCTCTCTTCTGACTTTTGAGCAAGGAGG 2439
Db 1999 CTCACCACAGAGCCCCCAGAGCCAAACAAGTCTCTCTCTTCTGACTTTTGAGCAAGGAGG 1940

Qy 2440 TCTGGATGTGGGAACCCCTTACCTCTGCTCTCTCTGGAGCCAGCCCTTCCAGAAACCCC 2499
Db 1939 TCTGGATGTGGGAACCCCTTACCTCTGCTCTCTCTGGAGCCAGCCCTTCCAGAAACCCC 1880

Qy 2500 AGCTCACCAGAGCGGAAGCAACCGTCCGAGCAGAGCACTGCAGCAGCTG 2550
Db 1879 AGCTCACCAGAGCGGAAGCAACCGTCCGAGCAGAGCACTGCAGCAGCTG 1829

RESULT 12
US-10-433-793-141
; Sequence 141, Application US/10433793
; Publication No. US2004012334A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von mit Angiogenese assoziierten Krankheiten
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/433,793
; CURRENT FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 212
; SEQ ID NO 141
; LENGTH: 6673
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-433-793-141

Query Match          6.3%; Score 291.6; DB 19; Length 6673;
Best Local Similarity 76.3%; Pred. No. 3.7e-76;
Matches 371; Conservative 0; Mismatches 114; Indels 1; Gaps 1;

Qy 5 GGGGAGCTGTCCTGTGGACGTGTAGCTTTCG-AGGTGACTCTGTACCATTCAGG 63
Db 5005 GGGGAGTGTGTGTGTGGAACGTGTAGTGTGTGAGAGGTGGATTTTGTATTATTAGG 5064

Qy 64 ATGTTTGAGGATGAGTATGTGTGCGCAGAGCACAATAAAGCAGCAGACCCCTTTGCC 123
Db 5065 ATGTTTGAGGATGAGTATGTGTGAGAGGTATATATAATAGTAGAGATTTTGTGTT 5124

Qy 124 CCTGCTTTCTCCCCCAACCAAGGCTGACCTGTGTCTTCTCCAGGTCTGGGATTTCTAAGT 183
```

```
Db 5125 TTGTGTTTTTTTTTTTAAATTAAGTGTGTTTTTTTTTAGGTTTGGGATTTTAAAGT 5184
Qy 184 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGACACAAGCCTGGGAGATGGCAGTGTATG 243
Db 5185 GATTTGTTTGTGTTTGGTCTCTCTCAGGATGAGTATAAGTTTGGGAGATGTTAGTGTATG 5244
Qy 244 GAAATGGCTGCCAGAGTGGCCCTCGCTCAGCAGTGGGGCAGCAGAGAACTCCCCCAAG 303
Db 5245 GAAATGGTGTGTTTGGTGTGTTTGGTGTAGTGGGGTGTAGTAGAAGGAAATTTTAA 5304
Qy 304 CCAAGGAGAGAGACGCCGACCTCGGGAAGAAACAGAGCTCCGTCTACAAGCTTGAGGCC 363
Db 5305 GTTAAGGAGAGAGAGCTGTTATTGTTGGGAAGAAATAGAGTTTCGTTTAAAGTTTGAAGTC 5364
Qy 364 GTGAGAGAGAGCCCTGTGTTCTCGGAAAGTGGGAGATCCCTGAATGAGCTGATTACCAAG 423
Db 5365 GTGAGAGAGAGTGTGTTGTTTTCGGAAGTGGGAGATTTTGAATGAGCTGATTATTAAAG 5424
Qy 424 GGCAAGCAAGAGAGGTCGAGGAGGGCCAGCTGCGCATCTCTATCATCGCCAGGCT 483
Db 5425 GGTATAGTTAAGGAAGGTTTCAGGTAGGGTTAGTTGTTATTTTATTATCGTTAGGTT 5484
Qy 484 GAGTGT 489
Db 5485 GAGTGT 5490

RESULT 13
US-09-918-995-2585
; Sequence 2585, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2585
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(499)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-2585

Query Match          5.8%; Score 267.4; DB 10; Length 499;
Best Local Similarity 95.6%; Pred. No. 2.8e-69;
Matches 307; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

Qy 2909 AGATCCCAGCTGCAGCCTTCAGCTTGTTCACCAAGAGCGGCGAGCCTGTTTCGTACGACA 2968
Db 182 AGATCCCAGCTGCAGCCTTCAGCTTGTTCACCAAGAGCGGCGAGCCTGTTTCGTACGACA 241
Qy 2969 TGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACATGGCCCCCTGATGGCAGCTTCG 3028
Db 242 TGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACATGGCCCCCTGATGGCAGCTTCG 301
Qy 3029 CTTGAGACTGGAGGTCAAGCATGGCCAGCTGGAGAAACAGGCCCTTAACCTGCCCTCCAC 3088
Db 302 CTTGAGACTGGAGGTCAAGCATGGCCAGCTGGAGAAACAGGCCCTTAACCTGCCCTCCAC 361
Qy 3089 CGCCGGCTCCACACTCGCCGAAAGACAGCCTTCTGCTCGGTGACGATGCTGCCCTGAAA 3148
Db 362 CGCCGGCTCCACACTCGCCG- AAGCAGCCTTCTGCTCGGTGACGATGCTGCCCTG-AA 419
```

QY	3149	ACACAGGCTGACGCGTTCCCAAGGGGATGCCAGCCCCCGCGCTCACAGTGGGAACCAAGG	3208
Db	420	ACACAGGCTGACGCGTTCCCAAGGGATGCCAGCCCCCGCGCTCACAGTGGGAACCAAGG	478
QY	3209	CCTCGCAGCAGCAAGGTGGGG	3229
Db	479	CCTCGCAGCAGCAAGGTGGAG	499

	Matches 267;	Conservative 0;	Mismatches 16;	Indels 4;	Gaps 4;
QY 3585	AAACCCCTTGCTGGCCTCACCTGCGCAGCTCACCCCTTTTGGGTGTAGGGGAAAAGAAT				3644
Db 10	AAACCCCTTGCTTGCCCTNACCTGCGCCAGCTCACCCCTTTTGGGTGTAGGGNAAAAGAAC				69
QY 3645	GCCTGACCTGGGAAGGCTCCCTGG-TAGAAATACACACACTTTTCAGGTTGTTCACAACA				3703
Db 70	GCCTGACCTGGGAAGGCTCCCTGGNTAGANTACCCACCTTTTAAGGTTGTTCACAACA				129
QY 3704	CAGGTCTCTGAGTTGACCTCTGGTTTCAGCCAAAGGACCAAGAGGTGTGTAAGTG-AAAGTG				3762
Db 130	NAGTCTCTGAGTTGACCTCTGGTTTCAGCCAAAGGACCAAGAGGTGTGTAAGTGAAAGTG				189
QY 3763	GTTCTCAGTCCCGACAGATGTGCCCTTTTG-CTGCTGGCTACCACTCTTCCCGCAGAGCAG				3821
Db 190	GTTCTNAGTCCCGACAGATGTGCCCTTTGGCTGTGGNTACCACTCTTCCCGCAGAGNAG				249
QY 3822	CAGGCCCGAGCCCC-TTCAGGCCGAGCACTGCCCGCAGACTCGCTGG				3867
Db 250	NAGGCCCGAGCCCCTTTNAAGGCCGAGNACTGCCCGCANATTTCGNTGG				296

Search completed: June 11, 2005, 06:13:18
 Job time : 1756.62 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:21:54 ; Search time 506.684 Seconds
(without alignments)
14842.245 Million cell updates/sec

Title: US-09-155-676B-6
Perfect score: 4596
Sequence: 1 agcgggggactgtgcgtg.....gatgacaatgttaaaaaa 4596

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 8181359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3149.6	68.5	3156	2	US-08-887-518-1
2	3149.6	68.5	3156	2	US-09-023-321-1
3	3149.6	68.5	3156	2	US-09-032-475-1
4	2837.6	61.7	2844	3	US-09-257-703-2
5	2837.6	61.7	2844	4	US-09-871-889A-2
6	288.2	6.5	362	4	US-09-513-999C-2583
7	88.8	1.9	2207	4	US-09-949-016-5071
8	88.8	1.9	2348	1	US-09-658-688A-3
9	83.8	1.8	3089	1	US-08-472-934-5
10	83.8	1.8	3089	2	US-08-323-460A-5
11	83.8	1.8	3089	2	US-08-461-146C-5
12	83.8	1.8	3089	3	US-08-461-145C-5
13	83.8	1.8	3089	3	US-08-628-829-9
14	83.8	1.8	3332	3	US-09-423-890-11
15	82.2	1.8	1935	3	US-09-423-890-5
16	61.6	1.3	3260	1	US-08-049-254-1
17	61.6	1.3	3260	1	US-08-472-934-1
18	61.6	1.3	3260	2	US-08-323-460A-1
19	61.6	1.3	3260	2	US-08-461-146C-1
20	61.6	1.3	3260	3	US-08-461-145C-1
21	61.6	1.3	3260	3	US-08-628-829-1
22	61.6	1.3	5253	3	US-09-423-890-7
23	61.6	1.3	5539	3	US-08-628-829-3
24	60.8	1.3	8563	4	US-09-902-540-3318
25	60.8	1.3	15351	4	US-09-902-540-1154
26	60.4	1.3	9862	4	US-09-691-861A-3
27	59.4	1.3	601	4	US-09-691-861A-14

28	58.8	1.3	1104	4	US-09-691-861A-1	Sequence 1, Appli
29	58.6	1.3	2169	4	US-09-016-434-1147	Sequence 1147, Ap
30	58.6	1.3	4162	2	US-08-459-448A-26	Sequence 26, Appl
31	58.6	1.3	4162	3	US-08-459-595A-26	Sequence 26, Appl
32	58.6	1.3	4162	3	US-08-459-504B-26	Sequence 26, Appl
33	58.6	1.3	4162	3	US-08-459-444-26	Sequence 26, Appl
34	58.6	1.3	4162	3	US-09-547-422-26	Sequence 26, Appl
35	58.6	1.3	4162	4	US-09-988-462-26	Sequence 26, Appl
36	58.6	1.3	4165	1	US-07-951-715A-26	Sequence 26, Appl
37	57.8	1.3	1732	3	US-09-430-564-1	Sequence 1, Appli
38	57.8	1.3	1816	4	US-09-762-258-1	Sequence 1, Appli
39	57.2	1.2	1458	4	US-09-230-896C-5	Sequence 5, Appli
40	57	1.2	320	4	US-09-270-767-536	Sequence 536, App
41	57	1.2	320	4	US-09-270-767-15818	Sequence 15818, A
42	56.4	1.2	1442	4	US-09-949-016-1590	Sequence 1590, Ap
43	56.4	1.2	1442	4	US-09-949-016-1591	Sequence 1591, Ap
44	56.4	1.2	1480	4	US-09-016-434-1454	Sequence 1454, Ap
45	55.2	1.2	1282	2	US-08-878-989-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-887-518-1
; Sequence 1, Application US/08887518
; Patent No. 5843721
; GENERAL INFORMATION:
; APPLICANT: Rothe, Mike
; APPLICANT: Wu, Lin
; TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/887,518
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3156 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-887-518-1

Query Match 68.5%; Score 3149.6; DB 2; Length 3156;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy	232	ATGCGAGTGTGGAATGGCTGCCAGGTGCCCTCGCTCAGAGTGGGCGACGAGAG 291
Db	1	ATGCGAGTGTGGAATGGCTGCCAGGTGCCCTCGCTCAGAGTGGGCGACGAGAG 60
Qy	292	GAACTCCCAAGCCAAAGAGAGACGCCCGCACTGGGGAAGAAACAGAGCTCGTCTAC 351

Db 61 GAACTCCCAAGACCAAGAGAGAGCGCGCCACTGGGGAAGAAACAGAGCTCCGTCTAC 120
QY 352 AAGCTTGAGGCGGTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGAGATCCTGAATGAC 411
Db 121 AAGCTTGAGGCGGTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGAGATCCTGAATGAC 180
QY 412 GTGATTACAGGGGACAGCCNAGGAAGCTCCGAGGCGAGGCCAGCTGCCATCTCTATC 471
Db 181 GTGATTACCAAGGGACAGCCNAGGAAGCTCCGAGGCGAGGCCAGCTGCCATCTCTATC 240
QY 472 ATGCGCCAGGCTGAGTGTGAGAAATAGCCAAGAGTTACGCCCCACCTTTTCAGAAAGCAATT 531
Db 241 ATGCGCCAGGCTGAGTGTGAGAAATAGCCAAGTTCAGGCCCACTTTTCAGAAAGCAATT 300
QY 532 TTCAATCGTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAT 591
Db 301 TTCAATCGTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAT 360
QY 592 GTGGCCCATGCTACAGAGGCAAAATGCCGCTGTGTTGGAAGGGAAGCGTCGCAGC 651
Db 361 GTGGCCCATGCTACAGAGGCAAAATGCCGCTGTGTTGGAAGGGAAGCGTCGCAGC 420
QY 652 AAAGCCCGAAGAAACGGAAGAAAGAGCTCAAAAGTCCCTGGCTCATGCGAGGTGGCC 711
Db 421 AAAGCCCGAAGAAACGGAAGAAAGAGCTCAAAAGTCCCTGGCTCATGCGAGGTGGCC 480
QY 712 TTGGCCAAACCCCTCCCAAGGACCCCTGAGCAGGAGAGTGCAATCCCAAGTGAGGAG 771
Db 481 TTGGCCAAACCCCTCCCAAGGACCCCTGAGCAGGAGAGTGCAATCCCAAGTGAGGAG 540
QY 772 GATGAGTCTCACTGGGCGCCCATATGTTAGAAACACCCGAGTTCACCAAGCCTCTG 831
Db 541 GATGAGTCTCACTGGGCGCCCATATGTTAGAAACACCCGAGTTCACCAAGCCTCTG 600
QY 832 AAGGAACAGGCTTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGCTTACGGCGGCT 891
Db 601 AAGGAACAGGCTTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGCTTACGGCGGCT 660
QY 892 CTGCTCTGATCAGAACTCCAAACATGATCAGCCCTTGCAATGTCTGAACCACTGTGG 951
Db 661 CTGCTCTGATCAGAACTCCAAACATGATCAGCCCTTGCAATGTCTGAACCACTGTGG 720
QY 952 AAATGCAACACCCAGAGCGAGGCCCTGCGCCCTGCCACGACCCCTTCCCTTA 1011
Db 721 AAATGCAACACCCAGAGCGAGGCCCTGCGCCCTGCCACGACCCCTTCCCTTA 780
QY 1012 AGCAGACTGCCTCATCCCTTCCCATTTCCACCCCTTCCAGCCCTTGAAACCTCACCCCTG 1071
Db 781 AGCAGACTGCCTCATCCCTTCCCATTTCCACCCCTTCCAGCCCTTGAAACCTCACCCCTG 840
QY 1072 GAGTCTTCTGGGCAAACTGGCCTGTGTAGACAGCAGAAACCTTGGCTGACCCACAC 1131
Db 841 GAGTCTTCTGGGCAAACTGGCCTGTGTAGACAGCAGAAACCTTGGCTGACCCACAC 900
QY 1132 CTGAGCAAACTGCGCTGTGTAGACAGTCCAAAGCCCTTGCCTGGCCCACTTGAGGCC 1191
Db 901 CTGAGCAAACTGCGCTGTGTAGACAGTCCAAAGCCCTTGCCTGGCCCACTTGAGGCC 960
QY 1192 AGCTGCTCTGCTGTGGTCCCATGAGAAATTTCTGTGGAGGAATACCTAGTGCATGCT 1251
Db 961 AGCTGCTCTGCTGTGGTCCCATGAGAAATTTCTGTGGAGGAATACCTAGTGCATGCT 1020
QY 1252 CTGCAAGGAGCGTGTAGCTCAAGCCAGGCCCAAGCTTGAACAGCTTGGCCAGACCTGG 1311
Db 1021 CTGCAAGGAGCGTGTAGCTCAAGCCAGGCCCAAGCTTGAACAGCTTGGCCAGACCTGG 1080
QY 1312 GCAGCAGGGGCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAACAGGCTGTC 1371
Db 1081 GCAGCAGGGGCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAACAGGCTGTC 1140
QY 1372 CTGCTCACTGAGAACTCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCC 1431

Db 1141 CTGCTCACTGAGAAACTCAAGCCAGTGGATTATGATGATACCGAGAAGATCCACTGGGCC 1200
QY 1432 ACGCACAGCTCCGCTGGCAGAGGCTCCTTCGAGAGGTGCACAGATGGAGCAAG 1491
Db 1201 ACGCACAGCTCCGCTGGCAGAGGCTCCTTCGAGAGGTGCACAGATGGAGCAAG 1260
QY 1492 CAGACTGGCTTCAGTGGCTGTCAAAAAGGTGGCTCGCTCGCTCGCTCGCTCGCTCGCT 1551
Db 1261 CAGACTGGCTTCAGTGGCTGTCAAAAAGGTGGCTCGCTCGCTCGCTCGCTCGCTCGCT 1320
QY 1552 CTGATGCAATGTGAGGATTTGACTCAACCAGAAATTTCTCCCTTTGTATGAGAGTGTGAGA 1611
Db 1321 CTGATGCAATGTGAGGATTTGACTCAACCAGAAATTTCTCCCTTTGTATGAGAGTGTGAGA 1380
QY 1612 GAAGGGCTTTGGGTCAACATCTTCATGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTG 1671
Db 1381 GAAGGGCTTTGGGTCAACATCTTCATGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTG 1440
QY 1672 GTCAAGGAGCAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGCCCTG 1731
Db 1441 GTCAAGGAGCAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGCCCTG 1500
QY 1732 GAGGCTCTGAATTAATCTCACTCAAGAAAGATTTCTGATGGGACGTCAAAAGCTGACAAC 1791
Db 1501 GAGGCTCTGAATTAATCTCACTCAAGAAAGATTTCTGATGGGACGTCAAAAGCTGACAAC 1560
QY 1792 GTGCTCTGTCCAGCGATGGAGCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1851
Db 1561 GTGCTCTGTCCAGCGATGGAGCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1620
QY 1852 CTTCAACCTGATGGCTTGGGAAAGTCTCTGTCTACAGGGGACTACATCCCTGGGCACAGAG 1911
Db 1621 CTTCAACCTGATGGCTTGGGAAAGTCTCTGTCTACAGGGGACTACATCCCTGGGCACAGAG 1680
QY 1912 ACCACATGGCTCCGAGGTGGTCTGCGGACGAGGTGCGACGCCCAAGGTGGATGTCTGG 1971
Db 1681 ACCACATGGCTCCGAGGTGGTCTGCGGACGAGGTGCGACGCCCAAGGTGGATGTCTGG 1740
QY 1972 AGCAGCTGCTATGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTCTCTTC 2031
Db 1741 AGCAGCTGCTATGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTCTCTTC 1800
QY 2032 CGAGGGCGCTCTGCCTCAAGATTGCCAGGAGCTCCGCTGTGAGGGAGATCCCAACC 2091
Db 1801 CGAGGGCGCTCTGCCTCAAGATTGCCAGGAGCTCCGCTGTGAGGGAGATCCCAACC 1860
QY 2092 TCCTGGGCCCTCTCAGGCCAGGCCNTCAAAGGGGCTGAGGAAAGGCCCATCCAC 2151
Db 1861 TCCTGGGCCCTCTCAGGCCAGGCCNTCAAAGGGGCTGAGGAAAGGCCCATCCAC 1920
QY 2152 CGGCTGCTGAGCGGAGCTGGGAGGGAAGGTGAAACGGGCACTACAGCAAGTGGGAGGT 2211
Db 1921 CGGCTGCTGAGCGGAGCTGGGAGGGAAGGTGAAACGGGCACTACAGCAAGTGGGAGGT 1980
QY 2212 CTGAAGAGCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCC 2271
Db 1981 CTGAAGAGCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCC 2040
QY 2272 AATTACACAGACCCCTCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCCAGGGCCC 2331
Db 2041 AATTACACAGACCCCTCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCCAGGGCCC 2100
QY 2332 CGGCCAGCTGAGGAGACAAACAGGAGAGCCCTTAAGCTCCAGCCCTCTCTCCACACAGAG 2391
Db 2101 CGGCCAGCTGAGGAGACAAACAGGAGAGCCCTTAAGCTCCAGCCCTCTCTCCACACAGAG 2160
QY 2392 CCCCCAGAGCCAAACAGTCTCTCTCTCTGATTTGAGCAAGGAGAGTCTGGGATGTGG 2451
Db 2161 CCCCCAGAGCCAAACAGTCTCTCTCTCTGACTTTGAGCAAGGAGAGTCTGGGATGTGG 2220
QY 2452 GAAACCTTACCTGTCTCTCTGAGGAGCCCTGCGAGAAACCCAGGCTCACCAGAG 2511
Db 2221 GAAACCTTACCTGTCTCTCTGAGGAGCCCTGCGAGAAACCCAGGCTCACCAGAG 2280

QY 772 GATGAGTCTCCACTCGGCGCCCATATGTTAGAAAACACCCCGCAGTTCACCAAGCCTCTG 831
DB 541 GATGAGTCTCCACTCGGCGCCCATATGTTAGAAAACACCCCGCAGTTCACCAAGCCTCTG 600
QY 832 AAGGAACAGAGCCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGGCGGCT 891
DB 601 AAGGAACAGAGCCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGGCGGCT 660
QY 892 CTGCTCTGATCAGAACTCCACAACTGATCAGCCCTTGCATGCTGAACACACGCTGG 951
DB 661 CTGCTCTGATCAGAACTCCACAACTGATCAGCCCTTGCATGCTGAACACACGCTGG 720
QY 952 AAATGACACACCCCGAGAGCGAGGCGCCCTGCCCCCTGCCACGACGCCCTTCCCTAT 1011
DB 721 AAATGACACACCCCGAGAGCGAGGCGCCCTGCCCCCTGCCACGACGCCCTTCCCTAT 780
QY 1012 AGCAGACTGCTCATCCCTTCCATTCACCTCTCCAGCCCTGGAACCTCAGCCTCTG 1071
DB 781 AGCAGACTGCTCATCCCTTCCATTCACCTCTCCAGCCCTGGAACCTCAGCCTCTG 840
QY 1072 GAGTCCTTCTGGGCAAACTGGCCTGTGTAGACAGCGAGAAACCTTGCCTGACCCACAC 1131
DB 841 GAGTCCTTCTGGGCAAACTGGCCTGTGTAGACAGCGAGAAACCTTGCCTGACCCACAC 900
QY 1132 CTGAGCAAACTGGCCTGTGTAGACAGTCGAAAGCCCTGCCCCCAGCCTGGAGCCC 1191
DB 901 CTGAGCAAACTGGCCTGTGTAGACAGTCGAAAGCCCTGCCCCCAGCCTGGAGCCC 960
QY 1192 AGCTGCTGTCTGCTGGTGCCTATGAGAAATTTCTGTGGAGGAATACCTAGTGATGCT 1251
DB 961 AGCTGCTGTCTGCTGGTGCCTATGAGAAATTTCTGTGGAGGAATACCTAGTGATGCT 1020
QY 1252 CTGCAAGGAGCGTGAGCTCAAGCCAGGCCCCAGCCTGACAGCCTGSCCAAGACCTGG 1311
DB 1021 CTGCAAGGAGCGTGAGCTCAAGCCAGGCCCCAGCCTGACAGCCTGSCCAAGACCTGG 1080
QY 1312 GCAGCAGCGGGTCCAGATCCCGGAGCCCCAGCCCCAACTGAGGACAAAGAGGTGTC 1371
DB 1081 GCAGCAGCGGGTCCAGATCCCGGAGCCCCAGCCCCAACTGAGGACAAAGAGGTGTC 1140
QY 1372 CTGCTCACTGAGAACTCAAGCAGTGGATATGATACCGAGAGAAAGTCCACTGGGCC 1431
DB 1141 CTGCTCACTGAGAACTCAAGCAGTGGATATGATACCGAGAGAAAGTCCACTGGGCC 1200
QY 1432 ACGCACAGCTCCGGCTGGGCAAGGCTCTTTCGAGAGGTCGACAGGATGGAGACAAG 1491
DB 1201 ACGCACAGCTCCGGCTGGGCAAGGCTCTTTCGAGAGGTCGACAGGATGGAGACAAG 1260
QY 1492 CAGACTGGCTTCCAGTGGCTGTCAAAAAGGTGCGCTGGAAAGTATTTGGGCGAGGAG 1551
DB 1261 CAGACTGGCTTCCAGTGGCTGTCAAAAAGGTGCGCTGGAAAGTATTTGGGCGAGGAG 1320
QY 1552 CTGATGGCATGTGAGGATGACCTCACCCAGAAATGTCCTTTGTATGGAGCTGTGAGA 1611
DB 1321 CTGATGGCATGTGAGGATGACCTCACCCAGAAATGTCCTTTGTATGGAGCTGTGAGA 1380
QY 1612 GAAGGGCTTGGGTCAACATCTTCATGGAGTGTCTGGAAGTGGCTCCCTGGGCGAGCTG 1671
DB 1381 GAAGGGCTTGGGTCAACATCTTCATGGAGTGTCTGGAAGTGGCTCCCTGGGCGAGCTG 1440
QY 1672 GTCAAGGAGCAGGCTGTCTCCAGAGCAGGCGCTGTACTACTCTGGGCGAGCCCTG 1731
DB 1441 GTCAAGGAGCAGGCTGTCTCCAGAGCAGGCGCTGTACTACTCTGGGCGAGCCCTG 1500
QY 1732 GAGGCTCTGGAACTCTCCACTCAGAAAGGATTCGTCATGGGAGCTCAAGCTGACAAAC 1791
DB 1501 GAGGCTCTGGAACTCTCCACTCAGAAAGGATTCGTCATGGGAGCTCAAGCTGACAAAC 1560
QY 1792 GTGCTCTTTCAGCGATGGGAGCCAGCCCTCTGTGACTTTGGGCCATGCTGTGTGT 1851
DB 1561 GTGCTCTTTCAGCGATGGGAGCCAGCCAGCCCTCTGTGACTTTGGGCCATGCTGTGTGT 1620
QY 1852 CTTCAACCTGATGGCCTGGGAAAGTCCTTGCTCACAGGGGACTACATCCCTGGGCACAG 1911

DB 1621 CTTCAACCTGATGGCCTGGGAAAGTCCTTGCTCACAGGGGACTACATCCCTGGGCACAG 1680
QY 1912 ACCACATGGCTCCGAGGTTGGTGTGGGCAAGAGCTGCGAGCGCCAAAGTGGATCTGTGG 1971
DB 1681 ACCACATGGCTCCGAGGTTGGTGTGGGCAAGAGCTGCGAGCGCCAAAGTGGATCTGTGG 1740
QY 1972 AGCAGCTCTGTATGATGCTGCATGCTCAACGGCTGCCACCCCTGACCTCAGTCTCTTC 2031
DB 1741 AGCAGCTCTGTATGATGCTGCATGCTCAACGGCTGCCACCCCTGACCTCAGTCTCTTC 1800
QY 2032 CGAGGGCGCTCTGCTTCAAGATTGCCAGGAGCCTTCGCTGTGAGGGAGATCCCAACC 2091
DB 1801 CGAGGGCGCTCTGCTTCAAGATTGCCAGGAGCCTTCGCTGTGAGGGAGATCCCAACC 1860
QY 2092 TCCTGGCGCCCTCTCACAGCCAGGCCATCAAGAGGGGCTGAGGAAAGAGCCATCCAC 2151
DB 1861 TCCTGGCGCCCTCTCACAGCCAGGCCATCAAGAGGGGCTGAGGAAAGAGCCATCCAC 1920
QY 2152 CGCGTGTCTGCAGCGGAGCTGGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGAGGT 2211
DB 1921 CGCGTGTCTGCAGCGGAGCTGGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGAGGT 1980
QY 2212 CTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGATCCACCGCCCAATCAAGCC 2271
DB 1981 CTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGATCCACCGCCCAATCAAGCC 2040
QY 2272 AATTACACACAGCCCTCATGCCAGCCGAGAGAGCTTTGCCNAGGGCCCCAGGGCCC 2331
DB 2041 AATTACACACAGCCCTCATGCCAGCCGAGAGAGCTTTGCCNAGGGCCCCAGGGCCC 2100
QY 2332 CGGCAGCTGAGGAGACAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACCACAG 2391
DB 2101 CGGCAGCTGAGGAGACAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCTCCACCACAG 2160
QY 2392 CCCCAGAGCCAAACAAGTCTCTCTCTTGAATTTGAGAAAGGAGTCTGGGATGTGG 2451
DB 2161 CCCCAGAGCCAAACAAGTCTCTCTCTTGAATTTGAGAAAGGAGTCTGGGATGTGG 2220
QY 2452 GAACCTTTACCTCTGCTCTCTGAGCGAGCCCTTGCAGAGAAACCCGAGCTCACCAGAG 2511
DB 2221 GAACCTTTACCTCTGCTCTCTGAGCGAGCCCTTGCAGAGAAACCCGAGCTCACCAGAG 2280
QY 2512 CGGAAAGCAACCGTCCCGAGCAGAACTGCAGCAGCTGGAAATAGAAATTTCTCTCAAC 2571
DB 2281 CGGAAAGCAACCGTCCCGAGCAGAACTGCAGCAGCTGGAAATAGAAATTTCTCTCAAC 2340
QY 2572 AGCTGTCCCAGCCATTTTCTCTGAGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATC 2631
DB 2341 AGCTGTCCCAGCCATTTTCTCTGAGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATC 2400
QY 2632 GACAGCTCTCTCTGCTGGATGACAGTGAAGAACCCATCAAGGCTCTCAAGGCTCG 2691
DB 2401 GACAGCTCTCTCTGCTGGATGACAGTGAAGAACCCATCAAGGCTCTCAAGGCTCG 2460
QY 2692 CGGACACCTCTGAGCTCAGGCTGACCTCTGAGAGCAGCCAGGCGCTCGAAGCTCC 2751
DB 2461 CGGACACCTCTGAGCTCAGGCTGACCTCTGAGAGCAGCCAGGCGCTCGAAGCTCC 2520
QY 2752 AGCTGGAACATGTGTGGCCCCGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGT 2811
DB 2521 AGCTGGAACATGTGTGGCCCCGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGT 2580
QY 2812 GTGAAAGTCCAAATACAGTCTCTTAATGTTGAACACCTGCACATCCGGGAGTTCACCGG 2871
DB 2581 GTGAAAGTCCAAATACAGTCTCTTAATGTTGAACACCTGCACATCCGGGAGTTCACCGG 2640
QY 2872 GTCAAAGTGGGAGACATCGCCACTCGGCATCAGCAGCCAGATCCAGCTGCAAGCTTCAGC 2931
DB 2641 GTCAAAGTGGGAGACATCGCCACTCGGCATCAGCAGCCAGATCCAGCTGCAAGCTTCAGC 2700
QY 2932 TTGCTCACCAAGACGGGCGAGCCTGTTCTGCTACACATGGAGGTGCCAGACTCGGGCATC 2991

Db 2701 TTGCTACCAAGACGGGCGCCCTGTTGGCTACGACATGAGAGTGCAGACTCGGGCATC 2760
Qy 2992 GACCTGCAGTGACACACTGGCCCTTGATGGCAGCTTCGCCCTGGAGCTGGAGGGTCAAGCAT 3051
Db 2761 GACCTGCAGTGACACACTGGCCCTTGATGGCAGCTTCGCCCTGGAGCTGGAGGGTCAAGCAT 2820
Qy 3052 GGCAGCTGGAGAACAGCCCTTAACCTTGCCTCCACGGCCGGCTCCACACTGCGCGGAAA 3111
Db 2821 GGCAGCTGGAGAACAGCCCTTAACCTTGCCTCCACGGCCGGCTCCACACTGCGCGGAAA 2880
Qy 3112 GCAGCCTTCTCTGCTCGGTGCACAGTGTGCCCTGAAACACAGGCTCAGCGGTTCCCGAGG 3171
Db 2881 GCAGCCTTCTCTGCTCGGTGCACAGTGTGCCCTGAAACACAGGCTCAGCGGTTCCCGAGG 2940
Qy 3172 GGATTGCGAGCCCCCGCTCACAGTGGGAACACAGGCGCTCGCAGCAGCAAGTGGGGGC 3231
Db 2941 GGATTGCGAGCCCCCGCTCACAGTGGGAACACAGGCGCTCGCAGCAGCAAGTGGGGGC 3000
Qy 3232 AAGCAGATGCTCCAGGATTTACACCTGAGCCCTGCGCCACCTGCTGAAAAACAT 3291
Db 3001 AAGCAGATGCTCCAGGATTTACACCTGAGCCCTGCGCCACCTGCTGAAAAACAT 3060
Qy 3292 CCGCCAGCTGAAGAGACAGAGGAGGATGGCAGGATTTACCTGGGAAACAAAAACAGGGA 3351
Db 3061 CCGCCAGCTGAAGAGACAGAGGAGGATGGCAGGATTTACCTGGGAAACAAAAACAGGGA 3120
Qy 3352 TCTTTTCTGCCCCCTGCTCCAGTGCAGTGGGCTGA 3387
Db 3121 TCTTTTCTGCCCCCTGCTCCAGTGCAGTGGGCTGA 3156

RESULT 3

US-09-032-475-1
; Sequence 1, Application US/09032475
; Patent No. 5854003
; GENERAL INFORMATION:
; APPLICANT: Rothe, Mike
; APPLICANT: Wu, Lin
; TITLE OF INVENTION: NFK Proteins, Nucleic Acids and Methods
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,475
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/887,518
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36, 627
; REFERENCE/DOCKET NUMBER: T97-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3156 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna

US-09-032-475-1
Query Match 68.5%; Score 3149.6; DB 2; Length 3156;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 232 ATGGCAGTGTGATAATGGCTTGCCTCCAGGTGCCCCCTGGCTCAGCAGTGGGGCAGCAGAG 291
Db 1 ATGGCAGTGTGATAATGGCTTGCCTCCAGGTGCCCCCTGGCTCAGCAGTGGGGCAGCAGAG 60
Qy 292 GAATCTCCCAAGCCAAAGAGAGAGAGCCGCCACTGGGGAGAGAAACAGAGCTCCGCTCTAC 351
Db 61 GAATCTCCCAAGCCAAAGAGAGAGAGCCGCCACTGGGGAGAGAAACAGAGCTCCGCTCTAC 120
Qy 352 AAGCTTGAGCCCGTGGAGAGAGAGCCCTGTGTCTTCGCGGAAGTGGGAGATCCTGAAATGAC 411
Db 121 AAGCTTGAGCCCGTGGAGAGAGAGCCCTGTGTCTTCGCGGAAGTGGGAGATCCTGAAATGAC 180
Qy 412 GTGATTACCAAGGGCAGCAGCAAGGAAGGCTCCGAGGCAGGGCCAGCTGCATCTCTATC 471
Db 181 GTGATTACCAAGGGCAGCAGCAAGGAAGGCTCCGAGGCAGGGCCAGCTGCATCTCTATC 240
Qy 472 ATCGCCAGGCTCAGTGTGAGATAGCACAAGATTTAGCCCCCAGCTTTTCAGAAACGATTT 531
Db 241 ATCGCCAGGCTCAGTGTGAGATAGCACAAGATTTAGCCCCCAGCTTTTCAGAAACGATTT 300
Qy 532 TTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAT 591
Db 301 TTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAT 360
Qy 592 GTGGCCCATGCTACAGAGGGCAAAATGGCCCGTGTGTGTGGAGGGAAGCGTCGGAGC 651
Db 361 GTGGCCCATGCTACAGAGGGCAAAATGGCCCGTGTGTGTGGAGGGAAGCGTCGGAGC 420
Qy 652 AAAGCCCGGAAGAAACGGAAGAGAGAGTCAAAGTCCCTGCTCATGAGGAGTGGCC 711
Db 421 AAAGCCCGGAAGAAACGGAAGAGAGAGTCAAAGTCCCTGCTCATGAGGAGTGGCC 480
Qy 712 TTGGCCAAACCCCTCCCGAGGACCCCTGAGCAGGAGAGTGCACCATCCAGTGCAGGAG 771
Db 481 TTGGCCAAACCCCTCCCGAGGACCCCTGAGCAGGAGAGTGCACCATCCAGTGCAGGAG 540
Qy 772 GATGAGTCTCCACTCGGGCGCCCATATGTTAGAAAAACCCCGCAGTTACCAAGCCCTGTG 831
Db 541 GATGAGTCTCCACTCGGGCGCCCATATGTTAGAAAAACCCCGCAGTTACCAAGCCCTGTG 600
Qy 832 AAGGAACCCAGGCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGCCCGGCT 891
Db 601 AAGGAACCCAGGCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGCCCGGCT 660
Qy 892 CTGCCTCGATCAGAACTCCCAAACTGATCAGCCCTTGCATGCTGAACCAAGCTGTGG 951
Db 661 CTGCCTCGATCAGAACTCCCAAACTGATCAGCCCTTGCATGCTGAACCAAGCTGTGG 720
Qy 952 AAATGCAACACCCCGCAGGAGCGGCGCCCTGCGCCCTGCGCCACGCAACCCCTTCCCTAT 1011
Db 721 AAATGCAACACCCCGCAGGAGCGGCGCCCTGCGCCCTGCGCCACGCAACCCCTTCCCTAT 780
Qy 1012 AGCAGACTGCTCATCCCTTCCCACTTCCAGCCCTTCCAGCCCTTCCAGCCCTTCCAGCCCTG 1071
Db 781 AGCAGACTGCTCATCCCTTCCCACTTCCAGCCCTTCCAGCCCTTCCAGCCCTTCCAGCCCTG 840
Qy 1072 GAGTCTCTCTGGGCAAACTGGGCTGTGTAGACAGCCAGCAAAACCCCTTGCCTGACCCACAC 1131
Db 841 GAGTCTCTCTGGGCAAACTGGGCTGTGTAGACAGCCAGCAAAACCCCTTGCCTGACCCACAC 900
Qy 1132 CTGAGCAAACTGGCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTCTGGAGGCC 1191
Db 901 CTGAGCAAACTGGCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTCTGGAGGCC 960
Qy 1192 AGCTGCCCTGTCTCGTGGTGGCCCATGAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCT 1251
Db 961 AGCTGCCCTGTCTCGTGGTGGCCCATGAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCT 1020

QY	1252	CTGCAAGGCAGCGTGAGCTCAAGCCAGGCCACAGCCCTGACCAAGACTGGCCCAAGACTGG	1311
DB	1021	CTGCNAGGCAGCGTGAGCTCAAGCCAGGCCACAGCCCTGACCAAGCTGGCCCAAGACTGG	1080
QY	1312	GCAGCAGCGGGCTCCAGATCCCGGGAGCCCAAGCCCAAACTGAGGACAACGAGGGGTGC	1371
DB	1081	GCAGCAGCGGGCTCCAGATCCCGGGAGCCCAAGCCCAAACTGAGGACAACGAGGGGTGC	1140
QY	1372	CTGCTCACTGAGAACTCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCC	1431
DB	1141	CTGCTCACTGAGAACTCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCC	1200
QY	1432	ACGCACAGCTCCGCTGGGAGAGGCTCTTTGGAGAGTGCACAGATGGAGGACAAG	1491
DB	1201	ACGCACAGCTCCGCTGGGAGAGGCTCTTTGGAGAGTGCACAGATGGAGGACAAG	1260
QY	1492	CAGACTGGCTTCCAGTGGCTGTCAAAAAGGTGGCTTGAAGTATTTCGGGCAGAGGAG	1551
DB	1261	CAGACTGGCTTCCAGTGGCTGTCAAAAAGGTGGCTTGAAGTATTTCGGGCAGAGGAG	1320
QY	1552	CTGATGGCATGTGCAGGATTGACCTACCAGAAATTGCCCTTTGTTATGAGCTGTGAGA	1611
DB	1321	CTGATGGCATGTGCAGGATTGACCTACCAGAAATTGCCCTTTGTTATGAGCTGTGAGA	1380
QY	1612	GAAGGCTTGGGTCAACATCTTCATCGAGCTGTGAAAGTGGCTCCCTGGGCCACAGCTG	1671
DB	1381	GAAGGCTTGGGTCAACATCTTCATCGAGCTGTGAAAGTGGCTCCCTGGGCCACAGCTG	1440
QY	1672	GTCGAAGGAGCAGGGCTGTCTCCACAGAGACCGGGCCCTGTACTACCTGGGCCCAGGCCCTG	1731
DB	1441	GTCGAAGGAGCAGGGCTGTCTCCACAGAGACCGGGCCCTGTACTACCTGGGCCCAGGCCCTG	1500
QY	1732	GAGGGTCTGGAACTCCTCCTCACTACGAAAGATTCTGCATGGGACGTCAAAGCTGACAAC	1791
DB	1501	GAGGGTCTGGAACTCCTCCTCACTACGAAAGATTCTGCATGGGACGTCAAAGCTGACAAC	1560
QY	1792	GTGCTCTGTCTCAGCGATGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGTGT	1851
DB	1561	GTGCTCTGTCTCAGCGATGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGTGT	1620
QY	1852	CTTCAACCTGATGGCCCTGGGAAAGTCTTGCTCACAGGGGACTACATCCCTGGCCACAGAG	1911
DB	1621	CTTCAACCTGATGGCCCTGGGAAAGTCTTGCTCACAGGGGACTACATCCCTGGCCACAGAG	1680
QY	1912	ACCCATATGGCTCCGAGGTGGTCTGGGCAGGAGTGCACGCAAGTGGATGCTGG	1971
DB	1681	ACCCATATGGCTCCGAGGTGGTCTGGGCAGGAGTGCACGCAAGTGGATGCTGG	1740
QY	1972	AGCAGCTGCTGATGATGCTGCAATGCTCAAAGGCTGCCACCCCTGGACTAGTTCCTTC	2031
DB	1741	AGCAGCTGCTGATGATGCTGCAATGCTCAAAGGCTGCCACCCCTGGACTAGTTCCTTC	1800
QY	2032	CGAGGGCCGCTTGCCTCAAGATTGGCAGCGAGCCTCCGCTCTGAGGGAGATGCCACCC	2091
DB	1801	CGAGGGCCGCTTGCCTCAAGATTGGCAGCGAGCCTCCGCTCTGAGGGAGATGCCACCC	1860
QY	2092	TCCTGCCCCCTCTCAAGCCCAAGCCATCCAAAGGGGTGAGGAAAGAGCCATCCAC	2151
DB	1861	TCCTGCCCCCTCTCAAGCCCAAGCCATCCAAAGGGGTGAGGAAAGAGCCATCCAC	1920
QY	2152	CGGCTGTCTCAGCGGAGCTGGAGGGAGGTCGACCGGGCATACAGCAAGTGGGAGGT	2211
DB	1921	CGGCTGTCTCAGCGGAGCTGGAGGGAGGTCGACCGGGCATACAGCAAGTGGGAGGT	1980
QY	2212	CTGAAGAGCCCTTGGAGGGAGAAATAAAGAACCAAGAATCCACCGCCCAATCAAGCC	2271
DB	1981	CTGAAGAGCCCTTGGAGGGAGAAATAAAGAACCAAGAATCCACCGCCCAATCAAGCC	2040
QY	2272	AATTACCAACAGACCTTCATGCCCAAGCCGAGAGAGCTTTGCGCAAGGGCCCCAGGGCCC	2331
DB	2041	AATTACCAACAGACCTTCATGCCCAAGCCGAGAGAGCTTTGCGCAAGGGCCCCAGGGCCC	2100

Qy	2332	CGGCCAGCTGAGGAGACAACAGCAGAGAGCCCTTAAGCTCCAGCCTCTCTCCACACAGAG	2399	
Db	2101	CGGCCAGCTGAGGAGACAACAGCAGAGAGCCCTTAAGCTCCAGCCTCTCTCCACACAGAG	2160	
Qy	2392	CCCCAGAGCAAAACAAGTCTCTCTCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGG	2451	
Db	2161	CCCCAGAGCAAAACAAGTCTCTCTCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGG	2220	
Qy	2452	GAACCCCTTAAGCTCTGTCTCTCCCTGGAGCCAGCCCTTGCAGAGAAACCCAGCTCACAGAG	2511	
Db	2221	GAACCCCTTAAGCTCTGTCTCTCTCCCTGGAGCCAGCCCTTGCAGAGAAACCCAGCTCACAGAG	2280	
Qy	2512	CGGAAACCAACCTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAAATTTCTCTCAAC	2571	
Db	2281	CGGAAACCAACCTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAAATTTCTCTCAAC	2340	
Qy	2572	AGCCTGTCCAGGCATTTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTCCTTCAGCATC	2631	
Db	2341	AGCCTGTCCAGGCATTTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTCCTTCAGCATC	2400	
Qy	2632	GACAGCCTCTCCCTGTCCGATGACAGTGAAGAGAACCCATCAAGGCCTCTCAAGCTCG	2691	
Db	2401	GACAGCCTCTCCCTGTCCGATGACAGTGAAGAGAACCCATCAAGGCCTCTCAAGCTCG	2460	
Qy	2692	CGGGACACCCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCAGGCGAGGCTCGAAGCTCC	2751	
Db	2461	CGGGACACCCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCAGGCGAGGCTCGAAGCTCC	2520	
Qy	2752	AGCTGGAAATGTGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTG	2811	
Db	2521	AGCTGGAAATGTGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTG	2580	
Qy	2812	GTGAAGTCCAAATACAGTCTCTTAATGTGTGAACACTTGACACATCGGAGGTTCCACCGG	2871	
Db	2581	GTGAAGTCCAAATACAGTCTCTTAATGTGTGAACACTTGACACATCGGAGGTTCCACCGG	2640	
Qy	2872	GTCAAAAGTGGAGACATCGCCAATGGCATCAGCAGCCAGATCCCAAGCTGACAGCTTCAGC	2931	
Db	2641	GTCAAAAGTGGAGACATCGCCAATGGCATCAGCAGCCAGATCCCAAGCTGACAGCTTCAGC	2700	
Qy	2932	TTGGTCAACCAAGACGGGAGCCTGTTCCCTACGACATGGAGGTTGCCAGACTCGGGCATC	2991	
Db	2701	TTGGTCAACCAAGACGGGAGCCTGTTCCCTACGACATGGAGGTTGCCAGACTCGGGCATC	2760	
Qy	2992	GACCTGCAGTGCACATCGGCCCTCGATGGCAGCTTCGCCCTGGAGCTGGAGGTTCAAGCAT	3051	
Db	2761	GACCTGCAGTGCACATCGGCCCTCGATGGCAGCTTCGCCCTGGAGCTGGAGGTTCAAGCAT	2820	
Qy	3052	GGCCAGCTGAGAAACAGGCCCTTAAACCTGCCCTTCAACCGCCGGCTCCAATCTGCCGAAA	3111	
Db	2821	GGCCAGCTGAGAAACAGGCCCTTAAACCTGCCCTTCAACCGCCGGCTCCAATCTGCCGAAA	2880	
Qy	3112	GCAGCCTTCTGCTCGGTGCAAGATGCTCCCTTGAAAAACAAGGCTCAGCCGTTCCACGG	3171	
Db	2881	GCAGCCTTCTGCTCGGTGCAAGATGCTCCCTTGAAAAACAAGGCTCAGCCGTTCCACGG	2940	
Qy	3172	GGATTGCCAGCCCCCGGCTCAAGTGGAAACAGGGCTTCGACAGCAAGGTGGGGC	3231	
Db	2941	GGATTGCCAGCCCCCGGCTCAAGTGGAAACAGGGCTTCGACAGCAAGGTGGGGC	3000	
Qy	3232	AAGCAGAAATGCTCCACAGATTTCACTCTGAGCCTCCCAACCTGCTGAAAAACAT	3291	
Db	3001	AAGCAGAAATGCTCCACAGATTTCACTCTGAGCCTCCCAACCTGCTGAAAAACAT	3060	
Qy	3292	CCGCCACGTGAAGAGACAAGAGGAGATGGCAGGAGTTACCTGGGGAACAAAAACAGGA	3351	
Db	3061	CCGCCACGTGAAGAGACAAGAGGAGATGGCAGGAGTTACCTGGGGAACAAAAACAGGA	3120	
Qy	3352	TCCTTTTCTGCCCTGCTCAGTCAGATTTGGCCTGA	3387	
Db	3121	TCCTTTTCTGCCCTGCTCAGTCAGATTTGGCCTGA	3156	

RESULT 4									
US-09-257-703-2									
; Sequence 2, Application US/09257703									
; Patent No. 6265538									
; GENERAL INFORMATION:									
; APPLICANT: Greene, Warner C.									
; APPLICANT: Lin, Xin									
; APPLICANT: Gelezulinas, Romas									
; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED									
; TITLE OF INVENTION: BY TNF-ALPHA AND IL-1									
; FILE REFERENCE: 30448.61USU1									
; CURRENT APPLICATION NUMBER: US/09/257,703									
; CURRENT FILING DATE: 1999-02-25									
; EARLIER APPLICATION NUMBER: 60/076,299									
; EARLIER FILING DATE: 1998-02-27									
; NUMBER OF SEQ ID NOS: 2									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 2									
; LENGTH: 2844									
; TYPE: DNA									
; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)									
US-09-257-703-2									
Query Match 61.7%; Score 2837.6; DB 3; Length 2844;									
Best Local Similarity 99.9%; Pred. No. 0;									
Matches 2840; Conservative 0; Mismatches 4; Indels 0; Gaps 0;									
Qy	232	ATGGCAGTGATGGAAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAAG	291						
Db	1	ATGGCAGTGATGGAAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAAG	60						
Qy	292	GAATCCCAAGCCAAAGAGAGACGCCCACTGGGGAGAAACAGAGCTCCGCTCTAC	351						
Db	61	GAATCCCAAGCCAAAGAGAGACGCCCACTGGGGAGAAACAGAGCTCCGCTCTAC	120						
Qy	352	AAGCTTGAAGCGGTGGAGAAAGAGCCCTGTCTTCGCGAAAGTGGAGATCTGAATGAC	411						
Db	121	AAGCTTGAAGCGGTGGAGAAAGAGCCCTGTCTTCGCGAAAGTGGAGATCTGAATGAC	180						
Qy	412	GTGATTACCAAGGGCAGCAGCAAGAGAGCTCCGAGGAGCGGCCAGCTGCCATCTCTATC	471						
Db	181	GTGATTACCAAGGGCAGCAGCAAGAGAGCTCCGAGGAGCGGCCAGCTGCCATCTCTATC	240						
Qy	472	ATCGCCCAAGGCTGAGTGTGAGAATAGCAAGAGTTTCAGCCCACTTTCAGAAAGCATT	531						
Db	241	ATCGCCCAAGGCTGAGTGTGAGAATAGCAAGAGTTTCAGCCCACTTTCAGAAAGCATT	300						
Qy	532	TTATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAT	591						
Db	301	TTATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAT	360						
Qy	592	GTGGCCCATGCTACAGAGGGCAAAATGCGCTGTGTGCGAAGGGAAAGCGTCGGAGC	651						
Db	361	GTGGCCCATGCTACAGAGGGCAAAATGCGCTGTGTGCGAAGGGAAAGCGTCGGAGC	420						
Qy	652	AAAGCCCGGAAGAAACGGAAGAGAGCTCAAGTCCCTGGCTCATGCAAGGAGTGGCC	711						
Db	421	AAAGCCCGGAAGAAACGGAAGAGAGCTCAAGTCCCTGGCTCATGCAAGGAGTGGCC	480						
Qy	712	TTGGCCAAACCCCTCCCAAGACCCCTGAGCAGAGAGTGCACATCCCAAGTGGAGAG	771						
Db	481	TTGGCCAAACCCCTCCCAAGACCCCTGAGCAGAGAGTGCACATCCCAAGTGGAGAG	540						
Qy	772	GATGAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCCTTG	831						
Db	541	GATGAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCCTTG	600						
Qy	832	AAGGAACCAAGGCTTGGGCAACTCTGTTTTTAAGCAGCTTGGCAGGGGCTACGGCCGGCT	891						
Db	601	AAGGAACCAAGGCTTGGGCAACTCTGTTTTTAAGCAGCTTGGCAGGGGCTACGGCCGGCT	660						
Qy	892	CTGCTCGATCAGAACTCCCAAACTGATCAGCCCTTTGCAATGTCTGAAACCAAGTGTGG	951						

Db	661	CTGCTCGATCAGAACTCCCAAACTGATCAGCCCTTTGCAATGTCTGAAACCAAGTGTGG	720
Qy	952	AAACTGCACCAACCCCGAGGAGGCGCCCTGCGCTGCCCGAGCGACCCCTTCCCTTAT	1011
Db	721	AAACTGCACCAACCCCGAGGAGGCGCCCTGCGCTGCCCGAGCGACCCCTTCCCTTAT	780
Qy	1012	AGCAGCTGCTCATCCCTTCCCAATTCACCTCTCAGCCCTGGAAACCTCACCCTCTG	1071
Db	781	AGCAGCTGCTCATCCCTTCCCAATTCACCTCTCAGCCCTGGAAACCTCACCCTCTG	840
Qy	1072	GAGTCTTCTTGGGCAAACTGGCCCTGTGTAGACAGCAGCAAAACCTTGCCTGAGCCAC	1131
Db	841	GAGTCTTCTTGGGCAAACTGGCCCTGTGTAGACAGCAGCAAAACCTTGCCTGAGCCAC	900
Qy	1132	CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTGCCCTGCGCCACACTGAGAGCC	1191
Db	901	CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTGCCCTGCGCCACACTGAGAGCC	960
Qy	1192	AGTGCCTGTCTCGTGTGCCCATGAGAAGTTTCTGTGTGAGGAATACCTAGTGCATGCT	1251
Db	961	AGTGCCTGTCTCGTGTGCCCATGAGAAGTTTCTGTGTGAGGAATACCTAGTGCATGCT	1020
Qy	1252	CTGCAAGCGAGCGTGAGCTCAAGCCAGGCGCCACAGCCTGACCCAGCAAGACCTGG	1311
Db	1021	CTGCAAGCGAGCGTGAGCTCAAGCCAGGCGCCACAGCCTGACCCAGCAAGACCTGG	1080
Qy	1312	GCAGCAGGGGCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAAGAGGCTGTC	1371
Db	1081	GCAGCAGGGGCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAAGAGGCTGTC	1140
Qy	1372	CTGCTCACTGAGAACTCAAGCCAGTGGATTTAGTACCGAAGAAAGTCCACTGGGCC	1431
Db	1141	CTGCTCACTGAGAACTCAAGCCAGTGGATTTAGTACCGAAGAAAGTCCACTGGGCC	1200
Qy	1432	ACGACACAGCTCCGCTGGGCGAGGCTCTTCCGAGAGGTGCACAGATGGAGACAG	1491
Db	1201	ACGACACAGCTCCGCTGGGCGAGGCTCTTCCGAGAGGTGCACAGATGGAGACAG	1260
Qy	1492	CAGACTGGCTTCCAGTGGCTGTCAAAAAGGTGGCGCTTGAAGTATTTCCGGCGAGAG	1551
Db	1261	CAGACTGGCTTCCAGTGGCTGTCAAAAAGGTGGCGCTTGAAGTATTTCCGGCGAGAG	1320
Qy	1552	CTGATGCGATGTGAGGATTTGACCTCACCGAAGATTTGCTCTTTGTATGAGAGTGTGAG	1611
Db	1321	CTGATGCGATGTGAGGATTTGACCTCACCGAAGATTTGCTCTTTGTATGAGAGTGTGAG	1380
Qy	1612	GAAGGGCTTGGGTCAACATCTTCATGGAGCTGTGGAAGTGGCTCCCTGGGCCAGCTG	1671
Db	1381	GAAGGGCTTGGGTCAACATCTTCATGGAGCTGTGGAAGTGGCTCCCTGGGCCAGCTG	1440
Qy	1672	GTCAAGGAGCAGGGCTGTCTCCAGAGAGCCGGGCCCTGTACTACTTGGGCGAGGCCCTG	1731
Db	1441	GTCAAGGAGCAGGGCTGTCTCCAGAGAGCCGGGCCCTGTACTACTTGGGCGAGGCCCTG	1500
Qy	1732	GAGGGCTTGGAAATACCTCCACTCAAGAGGATTTCTGATGGGGAGCGTCAAAAGCTGACAAC	1791
Db	1501	GAGGGCTTGGAAATACCTCCACTCAAGAGGATTTCTGATGGGGAGCGTCAAAAGCTGACAAC	1560
Qy	1792	GTGCTCTGTCCAGCGATGGAGCCAGCGAGCCCTCTGTGACTTTGGCCATGCTGTGTGT	1851
Db	1561	GTGCTCTGTCCAGCGATGGAGCCAGCGAGCCCTCTGTGACTTTGGCCATGCTGTGTGT	1620
Qy	1852	CTTCAAGCTGATGCGCTGGGAAAGTCTTGTCTACAGGGAGTACTATCCCTGGGCACAGAG	1911
Db	1621	CTTCAAGCTGATGCGCTGGGAAAGTCTTGTCTACAGGGAGTACTATCCCTGGGCACAGAG	1680
Qy	1912	ACCCACATGGCTCGGAGGTGGTGTGGGCGAGGAGCTGCGACGCGCAAGAGTGGATGTCTGG	1971
Db	1681	ACCCACATGGCTCGGAGGTGGTGTGGGCGAGGAGCTGCGACGCGCAAGAGTGGATGTCTGG	1740
Qy	1972	AGCAGCTGTGTATGATGTGCATGCTCAACGGCTGCCACCCCTGAGACTCAGTCTTCTTC	2031
Db	1741	AGCAGCTGTGTATGATGTGCATGCTCAACGGCTGCCACCCCTGAGACTCAGTCTTCTTC	1800

2032 CGAGGGCCGCTCTGCTCAAGATTGTCAGCGAGCCTCCGCGTGTGAGGGAGATCCCAACC 2091
|
Db CGAGGGCCGCTCTGCTCAAGATTGTCAGCGAGCCTCCGCGTGTGAGGGAGATCCCAACC 1860
|
2092 TCCTGCGCCCTCTCAAGCCAGCCGATCCAAAGAGGGGTGAGAAAGAGCCATCCAC 2151
|
Db TCCTGCGCCCTCTCAAGCCAGCCGATCCAAAGAGGGGTGAGAAAGAGCCATCCAC 1920
|
2152 CGCGTCTGTCAGCGAGCTGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGT 2211
|
Db CGCGTCTGTCAGCGAGCTGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGT 1980
|
2212 CTGAAGAGCCCTTGGAGGGAGAAATATAAGAACCAAGACATCCACCGCAATCAAGCC 2271
|
Db CTGAAGAGCCCTTGGAGGGAGAAATATAAGAACCAAGACATCCACCGCAATCAAGCC 2040
|
2272 AATTACACAGACCTCTCATGCCCCAGCGAGAGCTTTTCGCAAGGGCCCCAGGGCCC 2331
|
Db AATTACACAGACCTCTCATGCCCCAGCGAGAGCTTTTCGCAAGGGCCCCAGGGCCC 2100
|
2332 CGSCCAGCTGAGGAGACACAGCAGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAG 2391
|
Db CGSCCAGCTGAGGAGACACAGCAGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAG 2160
|
2392 CCCCCAGAGCCAAAGTCTCTCCCTTTGACTTTGAGCAAGAGGAGTCTGGGATGTGG 2451
|
Db CCCCCAGAGCCAAAGTCTCTCCCTTTGACTTTGAGCAAGAGGAGTCTGGGATGTGG 2220
|
2452 GAAACCTTACCTGTGCTCCTTGAGCCAGCCCTGCGAGAAACCCAGCTCACAGAG 2511
|
Db GAAACCTTACCTGTGCTCCTTGAGCCAGCCCTGCGAGAAACCCAGCTCACAGAG 2280
|
2512 CGSAAAGCAACGTCGCGAGCAGGAACCTGCAGCAGCTGGAATAGAAATTTCTCAAC 2571
|
Db CGSAAAGCAACGTCGCGAGCAGGAACCTGCAGCAGCTGGAATAGAAATTTCTCAAC 2340
|
2572 AGCTGTCCAGCAATTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTCCCTCAGCATC 2631
|
Db AGCTGTCCAGCAATTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTCCCTCAGCATC 2400
|
2632 GACAGCCTCTCCTGTGCGATGACAGTGAAGAACCCTCAAGAGCCTCTCAAGCTCG 2691
|
Db GACAGCCTCTCCTGTGCGATGACAGTGAAGAACCCTCAAGAGCCTCTCAAGCTCG 2460
|
2692 CGGGACACCTGAGCTCAGGCGTACACTCTGAGCAGCCAGCCGAGGCTCGAAGCTCC 2751
|
Db CGGGACACCTGAGCTCAGGCGTACACTCTGAGCAGCCAGCCGAGGCTCGAAGCTCC 2520
|
2752 AGCTGGAAACATGCTGTCGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGGT 2811
|
Db AGCTGGAAACATGCTGTCGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGGT 2580
|
2812 GTGAAGTCCAATACAGTCTCTTAATGTTGAACACTTGCAATCCGGGAGTTCCACGG 2871
|
Db GTGAAGTCCAATACAGTCTCTTAATGTTGAACACTTGCAATCCGGGAGTTCCACGG 2640
|
2872 GTCAAAGTGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCAAGCTGCGCTTCAGC 2931
|
Db GTCAAAGTGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCAAGCTGCGCTTCAGC 2700
|
2932 TTGGTCACAAAGA CCGGAGCCTGTTGCTACGATGGAGGTGCCAGACTCGGGCATC 2991
|
Db TTGGTCACAAAGA CCGGAGCCTGTTGCTACGATGGAGGTGCCAGACTCGGGCATC 2760
|
2992 GACCTGCGATGCACACTGGCCCTGATGCGAGCTTTCGCTGGAGCTGAGGGTCAAGCAT 3051
|
Db GACCTGCGATGCACACTGGCCCTGATGCGAGCTTTCGCTGGAGCTGAGGGTCAAGCAT 2820
|
3052 GGCCAGCTGGAGAACAGGCCCTAA 3075
|
Db GGCCAGCTGGAGAACAGGCCCTAA 2844

RESULT 5

US-09-871-889A-2
; Sequence 2, Application US/09871889A
; Patent No. 6645728

GENERAL INFORMATION:

; APPLICANT: Greene, Warner C.
; APPLICANT: Lin, Xin
; APPLICANT: Gelezuinas, Romas
; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED BY TNF-ALP
; FILE REFERENCE: 30448.61USD1
; CURRENT APPLICATION NUMBER: US/09/871,889A
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 09/257,703
; PRIOR FILING DATE: 1999-02-25
; PRIOR APPLICATION NUMBER: 60/076,299
; PRIOR FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2844
; TYPE: DNA

; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)

US-09-871-889A-2

Query Match 61.7%; Score 2837.6; DB 4; Length 2844;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2840; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 232 ATGGCAGTGATGGAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGCGAGCAGAG 291
|
Db 1 ATGGCAGTGATGGAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGCGAGCAGAG 60
|
QY 292 GAATCCCAAAGCCAAAGGAGAGACGCGCCACTTGGGGAAGAAACAGAGCTCGTCTAC 351
|
Db 61 GAATCCCAAAGCCAAAGGAGAGACGCGCCACTTGGGGAAGAAACAGAGCTCGTCTAC 120
|
QY 352 AAGCTTGAGGCGGTGGAGAGAGCCCTGTGTTCTCGGAAAGTGGAGATCTGAATGAC 411
|
Db 121 AAGCTTGAGGCGGTGGAGAGAGCCCTGTGTTCTCGGAAAGTGGAGATCTGAATGAC 180
|
QY 412 GTGATTACCAAGGCGACAGCCAAAGAGGCTCCGAGGCGAGGCCAGCTCTCTATC 471
|
Db 181 GTGATTACCAAGGCGACAGCCAAAGAGGCTCCGAGGCGAGGCCAGCTCTCTATC 240
|
QY 472 ATCGCCCAAGCTGATGTGAGAAATAGCCAGAGTTCAAGCCCACTTTTCAGAACGCAAT 531
|
Db 241 ATCGCCCAAGCTGATGTGAGAAATAGCCAGAGTTCAAGCTTCAAGCCCACTTTTCAGAACGCAAT 300
|
QY 532 TTTCATCGTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAT 591
|
Db 301 TTTCATCGTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAT 360
|
QY 592 GTGGCCCATGCTACAGAGGGCAAAATGGCCCGTGTGTTGGAAAGGAAACGCTCGCAGC 651
|
Db 361 GTGGCCCATGCTACAGAGGGCAAAATGGCCCGTGTGTTGGAAAGGAAACGCTCGCAGC 420
|
QY 652 AAAGCCCGAAGAAACCGAAGAGAGTCAAGTCCCTGGCTCATGCAAGAGTGGCC 711
|
Db 421 AAAGCCCGAAGAAACCGAAGAGAGTCAAGTCCCTGGCTCATGCAAGAGTGGCC 480
|
QY 712 TTGGCCAAACCCCTCCAGGACCCCTGAGCAGAGAGTGCACCATCCAGTGCAGAG 771
|
Db 481 TTGGCCAAACCCCTCCAGGACCCCTGAGCAGAGAGTGCACCATCCAGTGCAGAG 540
|
QY 772 GATGAGTCTCACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCTCTG 831
|
Db 541 GATGAGTCTCACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCTCTG 600
|
QY 832 AAGGAAACAGGCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGCTACGCGCGCT 891
|
Db 601 AAGGAAACAGGCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGCTACGCGCGCT 660

Qy 892 CTGCTCGATCAGAACTCCAACTGATCAGCCCCCTTGCAATGTCTGAACCAAGTGTGG 951
Db 661 CTGCTCGATCAGAACTCCAACTGATCAGCCCCCTTGCAATGTCTGAACCAAGTGTGG 720
Qy 952 AAACTGCACACCCAGGAGCGAGGCCCCCTGCCCCCTGCCCCAGCCCTTCCCTAT 1011
Db 721 AAACTGCACACCCAGGAGCGAGGCCCCCTGCCCCCTGCCCCAGCCCTTCCCTAT 780
Qy 1012 AGCAGACTGCTCATCCCTTCCCATTCACACCTCTCAGACCCCTGGAACCTCACCCTCTG 1071
Db 781 AGCAGACTGCTCATCCCTTCCCATTCACACCTCTCAGACCCCTGGAACCTCACCCTCTG 840
Qy 1072 GAGTCTTCTGGGCAAACTGGCTGTGTAGACAGCCAGCAAACTCTTGCCTGACCCACAC 1131
Db 841 GAGTCTTCTGGGCAAACTGGCTGTGTAGACAGCCAGCAAACTCTTGCCTGACCCACAC 900
Qy 1132 CTGAGCAAACTGGGCTGTGTAGACAGTCCAAAGCCCCCTGCTGCCCCACACCTGGAGCCC 1191
Db 901 CTGAGCAAACTGGGCTGTGTAGACAGTCCAAAGCCCCCTGCTGCCCCACACCTGGAGCCC 960
Qy 1192 AGCTGCTGTCTCGTGTGGCCCATGAGAAATTTTCTGTGAGGAATACCTAGTGCATGCT 1251
Db 961 AGCTGCTGTCTCGTGTGGCCCATGAGAAATTTTCTGTGAGGAATACCTAGTGCATGCT 1020
Qy 1252 CTGCAAGCAGCGTGAAGCTCAAGCCAGGCCACAGCCCTGACAGCCCTGGCCAAAGCCTGG 1311
Db 1021 CTGCAAGCAGCGTGAAGCTCAAGCCAGGCCACAGCCCTGACAGCCCTGGCCAAAGCCTGG 1080
Qy 1312 GCAGCAGGGGCTCAGATCCCGGGAGCCAGCCCCCAAACTGAGGACAACAGAGGGTGTG 1371
Db 1081 GCAGCAGGGGCTCAGATCCCGGGAGCCAGCCCCCAAACTGAGGACAACAGAGGGTGTG 1140
Qy 1372 CTGCTCACTGAGAACTCAAGCCAGTGGATTTAGTACCGAGAAAGTCCACTGGGCC 1431
Db 1141 CTGCTCACTGAGAACTCAAGCCAGTGGATTTAGTACCGAGAAAGTCCACTGGGCC 1200
Qy 1432 ACGCACCAGCTCCGCTGGGAGAGGCTCTTCGGAGAGGTGCACAGGATGGAGGACAAG 1491
Db 1201 ACGCACCAGCTCCGCTGGGAGAGGCTCTTCGGAGAGGTGCACAGGATGGAGGACAAG 1260
Qy 1492 CAGACTGGCTTCCAGTGGCTGTCAAAAGTGGCTGGAAGTATTTCCGGGAGAGGAG 1551
Db 1261 CAGACTGGCTTCCAGTGGCTGTCAAAAGTGGCTGGAAGTATTTCCGGGAGAGGAG 1320
Qy 1552 CTGATGCACTGTGAGGATTCACCTCACCAGAAATTTGCTTGTATGGAGCTGTGAGA 1611
Db 1321 CTGATGCACTGTGAGGATTCACCTCACCAGAAATTTGCTTGTATGGAGCTGTGAGA 1380
Qy 1612 GAAGGGCTTGGGTCAACATCTTCATGGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1671
Db 1381 GAAGGGCTTGGGTCAACATCTTCATGGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1440
Qy 1672 GTCAGAGCAGGGTGTCTCCAGAGAGCCGGGCTGTACTACTCTGGGCCAGGCCCTG 1731
Db 1441 GTCAGAGCAGGGTGTCTCCAGAGAGCCGGGCTGTACTACTCTGGGCCAGGCCCTG 1500
Qy 1732 GAGGCTCTGGAATACCTCCACTCAGAGGATTTGCAATGGGAGCGTCAAGCTGCAAC 1791
Db 1501 GAGGCTCTGGAATACCTCCACTCAGAGGATTTGCAATGGGAGCGTCAAAAGCTGCAAC 1560
Qy 1792 GTGCTCTGTCCAGCGATGGAGCCAGCAGCCCTCTGTGTACTTTTGGCCATGTGTGTGT 1851
Db 1561 GTGCTCTGTCCAGCGATGGAGCCAAGCAGCCCTCTGTGTACTTTTGGCCATGTGTGTGT 1620
Qy 1852 CTTCAACCTGATGGCTGGGAAAGTCTTTGCTCAAGGGGACTACATCCCTGGCACAGAG 1911
Db 1621 CTTCAACCTGATGGCTGGGAAAGTCTTTGCTCAAGGGGACTACATCCCTGGCACAGAG 1680
Qy 1912 ACCCAGATGGCTCGGAGGTGGTCTGGGAGGAGCTGCGACGCAAGGTGGATGTCTGG 1971
Db 1681 ACCCAGATGGCTCGGAGGTGGTCTGGGAGGAGCTGCGACGCAAGGTGGATGTCTGG 1740

Qy 1972 AGCAGCTGTGTATGATGTGACATGCTCAACGGCTGCCACCCCTGGACTCAGTCTTTC 2031
Db 1741 AGCAGCTGTGTATGATGTGACATGCTCAACGGCTGCCACCCCTGGACTCAGTCTTTC 1800
Qy 2032 CGAGGGCCGCTCTGCCCTCAAGATTTGCCAGAGGCTCCGCTGTGAGGGAGATCCACCC 2091
Db 1801 CGAGGGCCGCTCTGCCCTCAAGATTTGCCAGAGGCTCCGCTGTGAGGGAGATCCACCC 1860
Qy 2092 TCTGTGCCCCCTCTCAGAGCCCAAGGCTCAACAGAGGGGCTGAGGAAAGAGCCCATCCAC 2151
Db 1861 TCTGTGCCCCCTCTCAGAGCCCAAGGCTCAACAGAGGGGCTGAGGAAAGAGCCCATCCAC 1920
Qy 2152 CGGTGTCTGAGCGGAGCTGGGAGGGAAGGTGAACCGGGGACTACAGCAAGTGGAGGT 2211
Db 1921 CGGTGTCTGAGCGGAGCTGGGAGGGAAGGTGAACCGGGGACTACAGCAAGTGGAGGT 1980
Qy 2212 CTGAAGAGCCCTTGGAGGGGAGAAATATAAAGAACCAAGACATCCACCGCAAAATCAAGCC 2271
Db 1981 CTGAAGAGCCCTTGGAGGGGAGAAATATAAAGAACCAAGACATCCACCGCAAAATCAAGCC 2040
Qy 2272 AATTACCAACAGACCCCTTCCATGCCAGCCGAGAGAGCTTTGCGCAAGGGCCCCAGGGCCC 2331
Db 2041 AATTACCAACAGACCCCTTCCATGCCAGCCGAGAGAGCTTTGCGCAAGGGCCCCAGGGCCC 2100
Qy 2332 CGGCAGCTGAGGAGACAACAGGAGAGCCCTTAAGCTCAGACCTCTCTCCCAACAGAG 2391
Db 2101 CGGCAGCTGAGGAGACAACAGGAGAGCCCTTAAGCTCAGACCTCTCTCCCAACAGAG 2160
Qy 2392 CCCCAGAGCCAAACAGTCTCTCCCTGACTTTGAGCAGGAGGAGTCTGGGATGTGG 2451
Db 2161 CCCCAGAGCCAAACAGTCTCTCTCTTGTGCTTGTGAGCAGGAGGAGTCTGGGATGTGG 2220
Qy 2452 GAACCCCTTACTCTGTCTCTCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACCAGAG 2511
Db 2221 GAACCCCTTACTCTGTCTCTCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACCAGAG 2280
Qy 2512 CGGAAAGCAACCGTCCCGAGCAGGAACTGCGAGCAGCTGGAATAAGATAATTTCTCTCAAC 2571
Db 2281 CGGAAAGCAACCGTCCCGAGCAGGAACTGCGAGCAGCTGGAATAAGATAATTTCTCTCAAC 2340
Qy 2572 AGCTGTCCCAGCCATTTCTCTGGAGGCGAGGACAAATTTCTCTGTGCTCAGCATC 2631
Db 2341 AGCTGTCCCAGCCATTTCTCTGGAGGCGAGGACAAATTTCTCTGTGCTCAGCATC 2400
Qy 2632 GACAGCTCTCTCTGTGCGATGACAGTGAAGAAACCCATCAAGAGCTCTCTCAAGCTCG 2691
Db 2401 GACAGCTCTCTCTGTGCGATGACAGTGAAGAAACCCATCAAGAGCTCTCTCAAGCTCG 2460
Qy 2692 CGGACACCTGTAGCTCAGGCTGACTCTCTGGAGCAGCAGGCGGAGGCTCGAAGTCC 2751
Db 2461 CGGACACCTGTAGCTCAGGCTGACTCTCTGGAGCAGCAGGCGGAGGCTCGAAGTCC 2520
Qy 2752 AGCTGGAACATGTGTGCGGCCCGGGGGGCGCCACCGACACCCCAAGCTATTTCAATGGT 2811
Db 2521 AGCTGGAACATGTGTGCGGCCCGGGGGGCGCCACCGACACCCCAAGCTATTTCAATGGT 2580
Qy 2812 GTGAAGTCCAAATACAGTCTCTTAATGTGGAACACCTGCACATCCGGAGTCTCCACCG 2871
Db 2581 GTGAAGTCCAAATACAGTCTCTTAATGTGGAACACCTGCACATCCGGAGTCTCCACCG 2640
Qy 2872 GTCAAGTGGGAGACATCGGCCACTGGCATCAGCAGCAGATCCAGCTGACGCTTCAGC 2931
Db 2641 GTCAAGTGGGAGACATCGGCCACTGGCATCAGCAGCAGATCCAGCTGACGCTTCAGC 2700
Qy 2932 TTGGTCAACAAAGCGGGAGCTGTTCGCTACGACATGAGGAGTCCAGACTCGGGCATC 2991
Db 2701 TTGGTCAACAAAGCGGGAGCTGTTCGCTACGACATGAGGAGTCCAGACTCGGGCATC 2760
Qy 2992 GACCTCAGTGCACACTGGGCCCTGTATGGCAGCTTCGCTGGAGCTGAGGGTCAAGCAT 3051
Db 2761 GACCTCAGTGCACACTGGGCCCTGTATGGCAGCTTCGCTGGAGCTGAGGGTCAAGCAT 2820
Qy 3052 GGCCAGCTGGAGAACAGGCCCTTAA 3075


```
|||||
Db 2821 GGCCAGCTGGAGAACAGGCCCTTAA 2844

RESULT 6
US-09-513-999C-2583
; Sequence 2583, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 2583
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 79..360
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 19
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 22
; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-2583

Query Match 6.5%; Score 298.2; DB 4; Length 362;
Best Local Similarity 95.9%; Pred. No. 3.2e-70;
Matches 306; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 194 GTGTTTGTTCTCTCAGGATGAGCACAAAGCCCTGGGAGATGGCAGTGATGGAAATGGCCT 253
Db 41 CAGATCGGCGCTTGGCCGATGAGCACAAAGCCCTGGGAGATGGCAGTGATGGAAATGGCCT 100
QY 254 GCCCAGGTCCCTGGCTCAGCAGTGGGCGACAGAGAACTCCCAAGCCAAAGGAGA 313
Db 101 GCCCAGGTCCCTGGCTCAGCAGTGGGCGACAGAGAACTCCCAAGCCAAAGGAGA 160
QY 314 AGACGCCGCCACTGGGGAGAAACAGAGCTCCGTCTACAGCTTCAGGCCCTGGAGAGA 373
Db 161 AGACGCCGCCACTGGGGAGAAACAGAGCTCCGTCTACAGCTTCAGGCCCTGGAGAGA 220
QY 374 GCCCTGTGTTCTGCGAAAGTGGGAGATCCTGAATGACGTGATTACCAAGGGCAGGCCA 433
Db 221 GCCCTGTGTTCTGCGAAAGTGGGAGATCCTGAATGACGTGATTACCAAGGGCAGGCCA 280
QY 434 AGGAAGGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 493
Db 281 AGGAAGGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 340
QY 494 ATAGCCAGAGGTTTCAGCCC 512
Db 341 ATAGCCAGAGGTTTCAGCCC 359

RESULT 7
US-09-949-016-5071
; Sequence 5071, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5071
; LENGTH: 2207
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5071

Query Match 1.9%; Score 88.8; DB 4; Length 2207;
Best Local Similarity 53.4%; Pred. No. 2.8e-13;
Matches 236; Conservative 0; Mismatches 197; Indels 9; Gaps 2;

QY 1625 TCACATCTTCATGGAGCTGCTGGAAAGTGGCTCCCTGGCCAGCTGGTCAAGGAGCAGG 1684
Db 1312 TGACCATCTTCATGGAGTACATGCCAGGGGCTCGTGAAGACACAGTTGAAGGCTTAGC 1371
QY 1685 GCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGCCCTGGAGGCTCTGGAAT 1744
Db 1372 GTGCTCTGACAGAGAGCGTGACCCGAAAGTACACGCGCAGATCCTGGAGGCGATGTCT 1431
QY 1745 ACCTCCACTCAGGAAGATTCTGCATGGGAGCTCAAGCTGACAACTGCTCTCTGTCCA 1804
Db 1432 ACCTGCACAGCAACATGATTGTTCACCGGGACATTAAGGGAGCCAAACATCTCCGAGACT 1491
QY 1805 GCGATGGGAGCCAGCAGCCCTCTGTGACTTTGGCCATCTGTGTCTTCAACTGATG 1864
Db 1492 CTGCTGGGA---ATGTAAGCTGGGGACTTTGGGCCAGCAAGCTGCGAGC----- 1543
QY 1865 GCCTGGGAAAGTCTTGTCTACAGGGACTATACCTCTGGCAGACAGACCACATGGCTC 1924
Db 1544 -GATCTGTATCTCGGGACCGGCGATGCGCTCCGTCACTGGCACCACCTACTGGATGAGCC 1602
QY 1925 CGGAGGTGCTCTGGCAGGAGCTCGCAGCCCAAGTGGATGCTGGAGCAGCTCTGTA 1984
Db 1603 CTGAGGTGATCAGCGCGCAGGGCTATGGAAGGAAGCAGCGTGTGGAGCTTGGCTGCA 1662
QY 1985 TGATGCTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCGAGGGCGCTCT 2044
Db 1663 CTGTGGTGGAGTCTGACAGAGAAACACCGTGGCAGAGTATGAAGCTATGAGCTATGCGCCGA 1722
QY 2045 GCCTCAAGATTGCCAGCGAGCC 2066
Db 1723 TCTTCAAGATTGCCACCAGCC 1744

RESULT 8
US-09-658-688A-3
; Sequence 3, Application US/09658688A
; Patent No. 6498035
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: William Gaarde
; APPLICANT: Brett P. Monia
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF MEK3 EXPRESSION
; FILE REFERENCE: RTS-0143
; CURRENT APPLICATION NUMBER: US/09/658,688A
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 88
; SEQ ID NO 3
; LENGTH: 2348
; TYPE: DNA
```



```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)...(1964)
US-09-658-688A-3

Query Match
Best Local Similarity 1.9%; Score 88.8; DB 4; Length 2348;
Matches 236; Conservative 0; Mismatches 197; Indels 9; Gaps 2;

Qy 1625 TCAACATCTTCATGGAGCTGCTGGAAGTGGCTCCCTGGGCGACGTGCTCAAGGAGCAGG 1684
Db 1399 TGACACATCTTCATGGAGTACATGCGAGGGGCTCGTGAAGACACAGTTGAGGCTTACG 1458

Qy 1685 GCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCGAGGCCCTGGAGGGTCTGGAAT 1744
Db 1459 GTGCTCTGACAGAGAGCGTGACCCGAAAGTACACGCGCAGATCCTGGAGGGCATGTCCT 1518

Qy 1745 ACCTCCACTCAGAGGATTTCTGATGGGAGCTCAAGAGCTGACAAGCTGCTCTGTCCA 1804
Db 1519 ACCTGACAGCAACATGATTTTACCGGGACATTAAGGGAGCCAAACATCCTCCGAGACT 1578

Qy 1805 GCGATGGGAGCACGCGACCCCTGTGTGACTTTGGCCATGCTGTGTCTTCAACCTGATG 1864
Db 1579 CTGCTGGGA---ATGTAAGCTGGGGACTTTGGGCGCAGCAAGCGCTGCAGAC----- 1630

Qy 1865 GCCTGGGAAAGTCTCTGTCTCAAGGGGACTACATCCCTGGGCGACAGACCCACATGGCTC 1924
Db 1631 -GATCTGTATGCTGGGAGCGGCGCTCGCTCCGTCACCTGCGCACACCCCTACTGGATGAGCC 1689

Qy 1925 CGGAGGTGGTCTGGGCGAGGAGCTGCGAGCCAGGTGGATGCTTGAGAGCAGCTGTGTA 1984
Db 1690 CTGAGGTGATCAGCGGCGAGGGCTATGGAAGAAAGCAGACGTGTGGAGCCTGGGGTGCA 1749

Qy 1985 TGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCGAGGGCGGCTCT 2044
Db 1750 CTGTGGTGGAGATGCTGACAGAGAAACACCGTGGGCGAGGATGTAAGAGCTATGCGCCCA 1809

Qy 2045 GCCTCAAGATTGCCAGCGAGCC 2066
Db 1810 TCTTCAAGATTGCCACCGAGCC 1831

RESULT 9
US-08-472-934-5
; Sequence 5, Application US/08472934
; Patent No. 5753446
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
; TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lalive and Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,934
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,421
; FILING DATE: 15-May-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/354,516
```


QY 1695 GCTGCTCCAGAGACCGGCGCTGTACTACTGGGCGAGCCCTGGAGGGCTGTGAAT 1744
Db 1775 GAGCTCTGACAGAGAGTGTGACCCGCAAGTACACCCCGCAGATTCTGGAGGGCATGTCA 1834
QY 1745 ACCTCCACTCAGGAAGATTCTGATGGGACGTCAGAAAGCTGACAACTGCTCTGTGCTCA 1804
Db 1835 ACCTGCACAGCAACATGATTGTGATCGGACATCAGGAGACCAATATCTCCGAGACT 1894
QY 1805 GCGATGGAGCCAGCGAGCCCTCTGTGACTTTGGCCATGTGTGTCTTCAACTGATG 1864
Db 1895 CAGCTGGGA---ATGTGAAGCTTGGGATTTGGGGCCAGCAACGCGCTACAGACCATCT 1951
QY 1865 GCCTGGGAAAGTCTCTGCTACAGGGGACTACATCCCTGGCAGACAGACCCACATGGCTC 1924
Db 1952 GCATGTGAGGACAGGCA-----TCGCTGTGCTGCTGCGACACCCCTACTTGGATGATC 2005
QY 1925 CGGAGGTGGTCTGCGGAGGAGCTGCGACGCCAAGGTGATGTCTGGAGCAGCTGCTGTA 1984
Db 2006 CTGAAGTCATCAGTGGCGAGGGCTATGGAAGAGGCGAGCTGTGAGCCCTGGGCTGTA 2065
QY 1985 TGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCGGCTCT 2044
Db 2066 CTGTGGTGGAAATGCTGACAGAGAAACCACTTGGGCGAGATATGAAGCTATGCTGCCA 2125
QY 2045 GCCTCAAGATTGCCAGCGAGCCTCC 2069
Db 2126 TTTTCAAGATTGCCAGCGCTAC 2150

RESULT 12

US-08-461-145C-5
; Sequence 5, Application US/08461145C
; Patent No. 6074861
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: NOVEL MEKK PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive and Cockfield, LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,145C
; FILING DATE: 5-JUNE-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/049,254
; FILING DATE: 11-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/323,460
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/11690
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04178
; FILING DATE: 15-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/354,516
; FILING DATE: 21-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kara, Catherine J.
; REGISTRATION NUMBER: P-41,106
; REFERENCE/DOCKET NUMBER: CPI-004CN1
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 400..2280
; US-08-461-145C-5

Query Match 1.8%; Score 83.8; DB 3; Length 3089;
Best Local Similarity 52.6%; Pred. No. 7.4e-12;
Matches 234; Conservative 0; Mismatches 202; Indels 9; Gaps 2;
QY 1625 TCAACATCTTCATGGAGCTGCTGGAAGGTGGCTCCCTGGGGCCAGCTGGTCAAGAGCAGG 1684
Db 1715 TCACCATCTTTATGGAGTATATGCCAGGGGCTCTGTAAAGACCCAGTTGAAGGCCCTACG 1774
QY 1685 GCTGTCTCCAGAGGACCGGCGCTGTACTACCTGGGCCAGGCCCTGGAGGGTCTGGAAT 1744
Db 1775 GAGCTCTGACAGAGAGTGTGACCCGGAAGTACACCCGGCAGATTCTGGAGGGCATGTCT 1834
QY 1745 ACCTCCACTCACGAAGGATTCTGCATGGGGAGCTGCAAGCTGACAAAGCTGCTCTGCTGCTCA 1804
Db 1835 ACCTGCACAGCAACATGATTGTGATCGGACATCAAGGGAGCCCAATATCTCCGAGACT 1894
QY 1805 GCGATGGAGCCAGCGAGCCCTCTGTGACTTTGGCCATGTGTGTCTTCAACTGATG 1864
Db 1895 CAGCTGGGA---ATGTGAAGCTTGGGATTTGGGGCCAGCAACGCGCTACAGACCATCT 1951
QY 1865 GCCTGGGAAAGTCTCTGCTCACAGGGGACTACATCCCTGGCAGACAGACCCACATGGCTC 1924
Db 1952 GCATGTGAGGACAGGCA-----TCGCTGTGCTGCTGCGACACCCCTACTTGGATGATC 2005
QY 1925 CGGAGGTGGTCTGCGGAGGAGCTGCGACGCCAAGGTGATGTCTGGAGCAGCTGCTGTA 1984
Db 2006 CTGAAGTCATCAGTGGCGAGGGCTATGGAAGAGGCGAGCTGTGGAGCCCTGGGCTGTA 2065
QY 1985 TGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCGGCTCT 2044
Db 2066 CTGTGGTGGAAATGCTGACAGAGAAACCACTTGGGCGAGATATGAAGCTATGCTGCCA 2125
QY 2045 GCCTCAAGATTGCCAGCGAGCCTCC 2069
Db 2126 TTTTCAAGATTGCCAGCGCTAC 2150

RESULT 13

US-08-628-829-9
; Sequence 9, Application US/08628829A
; Patent No. 633170
; GENERAL INFORMATION:
; APPLICANT: Johnson, Gary L.
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To External
; FILE REFERENCE: CPI-004DVP3
; CURRENT APPLICATION NUMBER: US/08/628,829A
; FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: 08/440,421
; EARLIER FILING DATE: 1995-05-15
; EARLIER APPLICATION NUMBER: 08/323,460
; EARLIER FILING DATE: 1994-10-14
; EARLIER APPLICATION NUMBER: 08/049,254
; EARLIER FILING DATE: 1993-05-15
; EARLIER APPLICATION NUMBER: 08/410,602
; EARLIER FILING DATE: 1995-04-24
; EARLIER APPLICATION NUMBER: 08/472,934
; EARLIER FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 9
; LENGTH: 3089
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (400)..(2277)
US-08-628-829-9

Query Match
Best Local Similarity 1.8%; Score 83.8; DB 3; Length 3089;
Matches 234; Conservative 0; Mismatches 202; Indels 9; Gaps 2;
QY 1625 TCAACATCTTCATGGAGCTGCTGGAAGGTGGCTCCCTGGCCAGCTGGTCAAGGACAGG 1684
Db 1715 TCACCATCTTTATGGAGTATATGCCAGGGGGCTCTGTAAGACCAAGTTGAAGGCCTACG 1774
QY 1685 GCTGTCTCCAGAGGACCGGGCCCTGTACTACTGGGGCAGGCCCTGGAGGGTCTGGAAT 1744
Db 1775 GAGCTCTGACAGAGAGTGTGACCCGCAAGTACACCCGGCAGATTTCTGGAGGGCATGTCT 1834
QY 1745 ACCTCCACTCACGAAGATTTGCAATGGGAGCTCAAGCTGACACGCTGCTCCTGTCCA 1804
Db 1835 ACCTGCACAGCAACATGATTTGCAATGGGAGCATCAAGGGAGCCATATCTCCGAGACT 1894
QY 1805 GCGATGGGAGCCACGAGCCCTCTGTGACTTTGGCCCATGCTGTGTCTTCAACCTGATG 1864
Db 1895 CAGCTGGGA--ATGTGAAGCTTGGGGATTTGGGGCCAGCAACGCCCTAGACCATCT 1951
QY 1865 GCCTGGGAAGTCTTGTCTACAGGGGATACATCCCTGGCCACAGAGACCCACATGGCTC 1924
Db 1952 GCATGTCAAGGACAGGCAT-----TCGCTCTGTCACTGGCACACCCCTACTGGATGATC 2005
QY 1925 CGGAGTGTGCTGGGACGAGCTGCGACGCAACACCTTGGGAGGACACGCTGTGGAGCTGTA 1984
Db 2006 CTGAAGTCAATCAGTGGCGAGGGCTATGGAAGAAAGGACACGCTGTGGAGCCTGGGCTGTA 2065
QY 1985 TGATGCTGCACATGCTCAACGGCTGCCACCCCTGGAGCTCAGTTCTTCCGAGGCGCTCT 2044
Db 2066 CTGTGTGGAATGCTGACAGAGAAACACCTTGGGACAGATGAGAGCTATGAGCTATGGCTGCA 2125
QY 2045 GCCTCAAGATTGCCAGGAGCCTCC 2069
Db 2126 TTTTCAAGATTGCCACCCAGCCTAC 2150

RESULT 14
US-09-423-890-11
; Sequence 11, Application US/09423890
; Patent No. 6312934
; GENERAL INFORMATION:
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEKK PROTEIN AND NUCLEIC ACID MOLECULES
; FILE REFERENCE: CPI-085CPPC
; CURRENT APPLICATION NUMBER: US/09/423,890
; CURRENT FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: USN 60/078,153
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: USN 60/099,165
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 3332
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (332)..(2209)
US-09-423-890-11

Query Match
Best Local Similarity 1.8%; Score 83.8; DB 3; Length 3332;
Matches 233; Conservative 0; Mismatches 203; Indels 9; Gaps 2;

Best Local Similarity 52.6%; Pred. No. 7.7e-12;
Matches 234; Conservative 0; Mismatches 202; Indels 9; Gaps 2;
QY 1625 TCAACATCTTCATGGAGCTGCTGGAAGGTGGCTCCCTGGCCAGCTGGTCAAGGACAGG 1684
Db 1647 TCACCATCTTTATGGAGTATATGCCAGGGGGCTCTGTAAGACCAAGTTGAAGGCCTACG 1706
QY 1685 GCTGTCTCCAGAGGACCGGGCCCTGTACTACTGGGGCAGGCCCTGGAGGGTCTGGAAT 1744
Db 1707 GAGCTCTGACAGAGAGTGTGACCCGCAAGTACACCCGGCAGATTTCTGGAGGGCATGTCT 1766
QY 1745 ACCTCCACTCACGAAGATTTGCAATGGGAGCTCAAGCTGACAAACGCTGCTCCTGTCCA 1804
Db 1767 ACCTGCACAGCAACATGATTTGCAATGGGAGCATCAAGGGAGCCATATCTCCGAGACT 1826
QY 1805 GCGATGGGAGCCACGAGCCCTCTGTGACTTTGGCCCATGCTGTGTCTTCAACCTGATG 1864
Db 1827 CAGCTGGGA--ATGTGAAGCTTGGGGATTTGGGGCCAGCAACGCCCTACAGACCATCT 1883
QY 1865 GCCTGGGAAGTCTTGTCTACAGGGGACTACATCCCTGGCCACAGAGACCCACATGGCTC 1924
Db 1884 GCATGTCAAGGACAGGCAT-----TCGCTCTGTCACTGGCACACCCCTACTGGATGATC 1937
QY 1925 CGGAGTGTGCTGGGACGAGCTGCGACGCAAGGTGATGTCTGAGAGCAGCTGCTGTA 1984
Db 1938 CTGAAGTCAATCAGTGGCGAGGGCTATGGAAGAAAGGACAGAGCTGTGGAGCCTGGGCTGTA 1997
QY 1985 TGATGCTGCACATGCTCAACGGCTGCCACCCCTGGAGCTCAGTTCTTCCGAGGCGCTCT 2044
Db 1998 CTGTGTGGAATGCTGACAGAGAAACACCTTGGGACAGATGAGAGCTATGAGCTATGGCTGCCA 2057
QY 2045 GCCTCAAGATTGCCAGGAGCCTCC 2069
Db 2058 TTTTCAAGATTGCCACCCAGCCTAC 2082

RESULT 15
US-09-423-890-5
; Sequence 5, Application US/09423890
; Patent No. 6312934
; GENERAL INFORMATION:
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEKK PROTEIN AND NUCLEIC ACID MOLECULES
; FILE REFERENCE: CPI-085CPPC
; CURRENT APPLICATION NUMBER: US/09/423,890
; CURRENT FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: USN 60/078,153
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: USN 60/099,165
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1935
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)..(1902)
US-09-423-890-5

Query Match
Best Local Similarity 1.8%; Score 82.2; DB 3; Length 1935;
Matches 233; Conservative 0; Mismatches 203; Indels 9; Gaps 2;
QY 1625 TCAACATCTTCATGGAGCTGCTGGAAGGTGGCTCCCTGGGGCAGCTGGTCAAGGAGCAGG 1684
Db 1340 TCACCATCTTTATGGAGTATATGCCAGGGGGCTCTGTAAGACCAAGTTGAAGGCCTACG 1399
QY 1685 GCTGTCTCCAGAGGACCGGGCCCTGTACTACTGGGGCCAGGCCCTGGAGGGTCTGGAAT 1744
Db 1400 GAGCTCTGACAGAGAGTGTGACCCGCAAGTACACCCGGCAGATTTCTGGAGGGCATGTCT 1459

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 06:13:19 ; Search time 11495.3 Seconds
(without alignments)
16321.674 Million cell updates/sec

Title: US-09-155-676B-6
Perfect score: 4596
Sequence: 1 agcggggggactgtgcctg.....gatgacaatgttaaaaaa 4596

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4554873 seqs, 20411521753 residues
Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCRTUS1_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCRTUS2_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/PCRTUS_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US092A_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US092B_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US093A_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US093B_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US098D_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US104A_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US104B_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US105A_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US105B_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US108A_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US108B_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US110_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
86: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
87: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
88: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
89: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
90: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
91: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
92: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq:*
93: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq:*
94: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
95: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
96: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
97: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
98: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
99: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
100: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
101: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
102: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
103: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
104: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
105: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
106: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
107: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
108: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
109: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
110: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
111: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*
112: /cgn2_6/ptodata/1/pna/US6042_COMB.seq:*
113: /cgn2_6/ptodata/1/pna/US6043_COMB.seq:*
114: /cgn2_6/ptodata/1/pna/US6044_COMB.seq:*
115: /cgn2_6/ptodata/1/pna/US6045_COMB.seq:*
116: /cgn2_6/ptodata/1/pna/US6046_COMB.seq:*

117: /cgn2_6/ptodata/1/pna/US6047_COMB.seq:
118: /cgn2_6/ptodata/1/pna/US6048_COMB.seq:
119: /cgn2_6/ptodata/1/pna/US6049_COMB.seq:
120: /cgn2_6/ptodata/1/pna/US6050_COMB.seq:
121: /cgn2_6/ptodata/1/pna/US6051_COMB.seq:
122: /cgn2_6/ptodata/1/pna/US6052_COMB.seq:
123: /cgn2_6/ptodata/1/pna/US6053_COMB.seq:
124: /cgn2_6/ptodata/1/pna/US6054_COMB.seq:
125: /cgn2_6/ptodata/1/pna/US6055_COMB.seq:
126: /cgn2_6/ptodata/1/pna/US6056_COMB.seq:
127: /cgn2_6/ptodata/1/pna/US6057_COMB.seq:
128: /cgn2_6/ptodata/1/pna/US6058_COMB.seq:
129: /cgn2_6/ptodata/1/pna/US6059_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4596	100.0	4596	17	US-09-155-676-6	Sequence 6, Appli
2	4596	100.0	4596	17	US-09-155-676A-6	Sequence 6, Appli
3	4578.4	99.6	4596	1	PCT-US01-08631-15294	Sequence 15294, A
4	4578.4	99.6	4596	1	PCT-US02-18947-975	Sequence 975, App
5	4578.4	99.6	4596	2	PCT-US04-01100-975	Sequence 975, App
6	4578.4	99.6	4596	16	US-09-053-375B-1234	Sequence 1234, Ap
7	4578.4	99.6	4596	41	US-09-949-003C-437	Sequence 437, App
8	4578.4	99.6	4596	45	US-09-981-397A-17	Sequence 17, Appl
9	4578.4	99.6	4596	49	US-10-172-118-975	Sequence 975, App
10	4578.4	99.6	4596	51	US-10-283-375A-294	Sequence 294, App
11	4578.4	99.6	4596	52	US-10-342-887-975	Sequence 975, App
12	4578.4	99.6	4596	117	US-60-474-850-144	Sequence 144, App
13	4440	96.6	4584	41	US-09-949-003C-1291	Sequence 1291, Ap
14	4440	96.6	4584	47	US-10-087-192-887	Sequence 887, App
15	4440	96.6	4684	49	US-10-170-235-9152	Sequence 9152, Ap
16	4439.6	96.6	4591	97	US-60-278-258-482	Sequence 482, Ap
17	4438	96.6	4591	102	US-60-324-185-32844	Sequence 32844, A
18	4437.6	96.6	4684	114	US-60-443-566-226	Sequence 226, App
19	4437.6	96.6	4684	115	US-60-453-050-5384	Sequence 5384, Ap
20	4437.6	96.6	4684	115	US-60-453-135-5384	Sequence 5384, Ap
21	4437.6	96.6	4684	115	US-60-455-444-2732	Sequence 2732, Ap
22	4437.6	96.6	4684	116	US-60-465-241-2732	Sequence 2732, Ap
23	4437.6	96.6	4684	116	US-60-466-412-5384	Sequence 5384, Ap
24	4437.6	96.6	4684	118	US-60-485-450-566	Sequence 566, App
25	4437.6	96.6	4684	118	US-60-487-610-912	Sequence 912, App
26	4437.6	96.6	4684	126	US-60-563-440-566	Sequence 566, App
27	4437.6	96.6	4684	128	US-60-582-609-912	Sequence 912, App
28	4423.8	96.3	4596	86	US-60-172-373-3494	Sequence 3494, Ap
29	4245.6	92.4	4458	49	US-10-170-235-10109	Sequence 10109, A
30	4243.2	92.3	4458	114	US-60-443-566-227	Sequence 227, App
31	4243.2	92.3	4458	115	US-60-453-050-5385	Sequence 5385, Ap
32	4243.2	92.3	4458	115	US-60-453-135-5385	Sequence 5385, Ap
33	4243.2	92.3	4458	115	US-60-455-444-2733	Sequence 2733, Ap
34	4243.2	92.3	4458	116	US-60-465-241-2733	Sequence 2733, Ap
35	4243.2	92.3	4458	116	US-60-466-412-5385	Sequence 5385, Ap
36	4243.2	92.3	4458	118	US-60-485-450-567	Sequence 567, App
37	4243.2	92.3	4458	118	US-60-487-610-913	Sequence 913, App
38	4243.2	92.3	4458	126	US-60-563-440-567	Sequence 567, App
39	4243.2	92.3	4458	128	US-60-582-609-913	Sequence 913, App
40	4211.8	91.6	4487	22	US-09-496-914A-6881	Sequence 6881, Ap
41	4211.8	91.6	4487	26	US-09-560-875A-6881	Sequence 6881, Ap
42	2831.4	61.6	2842	94	US-60-248-798-532	Sequence 532, App
43	2673.4	58.2	3152	62	US-10-760-320A-4571	Sequence 4571, Ap
44	2673.4	58.2	3152	62	US-10-760-620A-4571	Sequence 4571, Ap
45	2054.6	44.7	2829	48	US-10-144-771-10006	Sequence 10006, A

ALIGNMENTS

RESULT 1
US-09-155-676-6
; Sequence 6, Application US/09155676
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/155.676
; FILING DATE: 04-JAN-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IL97/00117
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117800
; FILING DATE: 02-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 119133
; FILING DATE: 26-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: WALLACH=21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4596 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-155-676-6
Query Match 100.0%; Score 4596; DB 17; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGCGGGGGGACGTGTCGGGTGGAACGTGTAGCTGTGAAGGTGGACTCTGTACCATTG 60
Db 1 AGCGGGGGGACGTGTCGGGTGGAACGTGTAGCTGTGAAGGTGGACTCTGTACCATTG 60
QY 61 AGGATGTTTGGAGGATGATGTGTGGCAGAGGCACATATAACAGGCAGAGACCCCTTT 120
Db 61 AGGATGTTTGGAGGATGATGTGTGGCAGAGGCACATATAACAGGCAGAGACCCCTTT 120
QY 121 GCCCTGCTCTTCTCCCAAGCTGACCTGTCTTCTCCAGGTCCTGGATTCTA 180
Db 121 GCCCTGCTCTTCTCCCAAGCTGACCTGTCTTCTCCAGGTCCTGGATTCTA 180
QY 181 AGTGACCTGCTCTGTGTTTGGTCTCTCTCAGATGAGCACAAGCCTGGAGATGGCAGTG 240
Db 181 AGTGACCTGCTCTGTGTTTGGTCTCTCTCAGATGAGCACAAGCCTGGAGATGGCAGTG 240

QY 241 ATGGAATGGCTGCGCCAGTGCCTGGCTCAGCAGTGGGCGAGGAAGAACTCCCC 300
DB 241 ATGGAATGGCTGCGCCAGTGCCTGGCTCAGCAGTGGGCGAGGAAGAACTCCCC 300
QY 301 AAGCCAAAGGAGAGACGCGCCACTGGGGAAGAAACAGAGCTCCGTCTTACAAGCTTGAG 360
DB 301 AAGCCAAAGGAGAGACGCGCCACTGGGGAAGAAACAGAGCTCCGTCTTACAAGCTTGAG 360
QY 361 GCCGTGGAGAGAGCCCTGTGTCTTCGGAAAGTGGAGATCCTGAATGA CGTGATTAAC 420
DB 361 GCCGTGGAGAGAGCCCTGTGTCTTCGGAAAGTGGAGATCCTGAATGA CGTGATTAAC 420
QY 421 AAGGSCACAGCCAAAGGAGCTCCAGCAGGCGCCAGCTGCCATCTCTATCATCGCCAG 480
DB 421 AAGGSCACAGCCAAAGGAGCTCCAGCAGGCGCCAGCTGCCATCTCTATCATCGCCAG 480
QY 481 GCTGAGTGTGAGAAATAGCAAGAGTTTCAAGCCCACTTTTCAAGACGCAATTTTCAATCGCT 540
DB 481 GCTGAGTGTGAGAAATAGCAAGAGTTTCAAGCCCACTTTTCAAGACGCAATTTTCAATCGCT 540
QY 541 GGGTCCAAACAGTACAGCCAGTCCAGAGTCTTGATCAGATCCCAACAAATGTGGCCCAT 600
DB 541 GGGTCCAAACAGTACAGCCAGTCCAGAGTCTTGATCAGATCCCAACAAATGTGGCCCAT 600
QY 601 GCTACAGAGGCGAAATGGCCGCTGTGTGGAAAGGAAAGCGTCCGAGCAAAAGCCCGG 660
DB 601 GCTACAGAGGCGAAATGGCCGCTGTGTGGAAAGGAAAGCGTCCGAGCAAAAGCCCGG 660
QY 661 AAGAAACGGAAGAGAGTCAAGTCCCTGGCTCATGAGGAGTGGCTTTGGCCAAA 720
DB 661 AAGAAACGGAAGAGAGTCAAGTCCCTGGCTCATGAGGAGTGGCTTTGGCCAAA 720
QY 721 CCCCTCCCAAGAGCCCTGAGCAGAGAGTGCACCATCCAGTGCAGGAGATGAGTCT 780
DB 721 CCCCTCCCAAGAGCCCTGAGCAGAGAGTGCACCATCCAGTGCAGGAGATGAGTCT 780
QY 781 CCATCGGCGCCCATATGTTAGAAACACCCGCGAGTTCAACAGCCCTCTGAAGAACCA 840
DB 781 CCATCGGCGCCCATATGTTAGAAACACCCGCGAGTTTCAACAGCCCTCTGAAGAACCA 840
QY 841 GGCTTGGGCAACTCTGTTTTAAGCAGCTTGGCAGGCGCTACGCGCGCTCTGCCTCGA 900
DB 841 GGCTTGGGCAACTCTGTTTTAAGCAGCTTGGCAGGCGCTACGCGCGCTCTGCCTCGA 900
QY 901 TCAGAACTCCCAAACTGATCAGCCCTTGCAATGTCTGAACCAAGTGTGGAACCTGCAC 960
DB 901 TCAGAACTCCCAAACTGATCAGCCCTTGCAATGTCTGAACCAAGTGTGGAACCTGCAC 960
QY 961 CACCCCAAGGAGGAGCCCTGCGCCCTGCGCCAGCAACCCCTTCCCTATAGCAGACTG 1020
DB 961 CACCCCAAGGAGGAGCCCTGCGCCCTGCGCCAGCAACCCCTTCCCTATAGCAGACTG 1020
QY 1021 CCTCATCCCTTCCCAATCCACCTCTCAGCCCTGGAAACCTCAACCTCTGGAGTCTTC 1080
DB 1021 CCTCATCCCTTCCCAATCCACCTCTCAGCCCTGGAAACCTCAACCTCTGGAGTCTTC 1080
QY 1081 CTGGCAAACTGGCTGTGTAGACAGCCAGAAACCTTTGCTGACCCACACCTGAGCAAA 1140
DB 1081 CTGGCAAACTGGCTGTGTAGACAGCCAGAAACCTTTGCTGACCCACACCTGAGCAAA 1140
QY 1141 CTGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGCCCAAGTGCCTG 1200
DB 1141 CTGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGCCCAAGTGCCTG 1200
QY 1201 TCTGCTGTGCGCCATGAGAAATTTCTGTGGAGGATACCTAGTGCATGCTCTGCAAGGC 1260
DB 1201 TCTGCTGTGCGCCATGAGAAATTTCTGTGGAGGATACCTAGTGCATGCTCTGCAAGGC 1260
QY 1261 AGCGTGAGCTCAAGCAGGCGCCACAGCTGTGACCCCTGGCCCAAGACCTGGGCGAGCAGG 1320
DB 1261 AGCGTGAGCTCAAGCAGGCGCCACAGCTGTGACCCCTGGCCCAAGACCTGGGCGAGCAGG 1320
QY 1321 GGCTCCAGATCCCGGAGCCAGCCCAAAACTGAGGCAACAGAGGCTGCTCTGCTCACT 1380

DB 1321 GGCTCCAGATCCCGGAGCCAGCCCAAAACTGAGGCAACAGAGGCTGCTCTGCTCACT 1380
QY 1381 GAGAACTCAAGCCAGTGGATTATAGTACCGAGAAGAGTCCACTGGGCGCACGACCCAG 1440
DB 1381 GAGAACTCAAGCCAGTGGATTATAGTACCGAGAAGAGTCCACTGGGCGCACGACCCAG 1440
QY 1441 CTCGCTGGGCGAGAGGCTCTCTCGGAGAGTGCACAGATGGAGGACAAAGCAGACTGCG 1500
DB 1441 CTCGCTGGGCGAGAGGCTCTCTCGGAGAGTGCACAGATGGAGGACAAAGCAGACTGCG 1500
QY 1501 TTCCAGTCCGCTGTCAAAAGAGTGCCTGGAAGTATTTCCGGCAGAGAGCTGATGGCA 1560
DB 1501 TTCCAGTCCGCTGTCAAAAGAGTGCCTGGAAGTATTTCCGGCAGAGAGCTGATGGCA 1560
QY 1561 TGTGAGGATTTGACCTCAACCCAGAAATTTCTCTTTGTATGGAGCTGTGAGAGAAGGCTCT 1620
DB 1561 TGTGAGGATTTGACCTCAACCCAGAAATTTCTCTTTGTATGGAGCTGTGAGAGAAGGCTCT 1620
QY 1621 TGGGTCAACATCTTCAATGGAGCTGTGAAAGTGGCTCCCTGGGCGAGCTGGTCAAGGAG 1680
DB 1621 TGGGTCAACATCTTCAATGGAGCTGTGAAAGTGGCTCCCTGGGCGAGCTGGTCAAGGAG 1680
QY 1681 CAGGCTGTCTCCAGAGGACCGGCGCTGTACTACTCTGGGCGAGGCTCTGGAGGCTGTG 1740
DB 1681 CAGGCTGTCTCCAGAGGACCGGCGCTGTACTACTCTGGGCGAGGCTCTGGAGGCTGTG 1740
QY 1741 GAATACCTCCACTCAACGAAGGATTTCTGATGGGAGCTCAAAAGCTGACAAAGTGTCTCTG 1800
DB 1741 GAATACCTCCACTCAACGAAGGATTTCTGATGGGAGCTCAAAAGCTGACAAAGTGTCTCTG 1800
QY 1801 TCCAGCGATGGAGCCACGCGAGCCCTCTGTGACTTTTGGCCATGTGTGTCTTCAACCT 1860
DB 1801 TCCAGCGATGGAGCCACGCGAGCCCTCTGTGACTTTTGGCCATGTGTGTCTTCAACCT 1860
QY 1861 GATGGCTGGGAAAGTCTTCTGCTCAGAGGAGCTACATCTCTGGGCGAGAGCCACATG 1920
DB 1861 GATGGCTGGGAAAGTCTTCTGCTCAGAGGAGCTACATCTCTGGGCGAGAGCCACATG 1920
QY 1921 GCTCGGAGTGGTGTCTGGGCGAGGAGTGCACGCGAAGTGGATGTCTGGAGAGTGC 1980
DB 1921 GCTCGGAGTGGTGTCTGGGCGAGGAGTGCACGCGAAGTGGATGTCTGGAGAGTGC 1980
QY 1981 TGTATGATGTGCACTGCTCAACGGCTGCCACCCCTGGA CTCACTTCTTCCGAGGCGG 2040
DB 1981 TGTATGATGTGCACTGCTCAACGGCTGCCACCCCTGGA CTCACTTCTTCCGAGGCGG 2040
QY 2041 CTCTGCTCAAGATTTGCCAGGAGCTCCGCTGTGAGGAGATCCCACTCTCTGCGCC 2100
DB 2041 CTCTGCTCAAGATTTGCCAGGAGCTCCGCTGTGAGGAGATCCCACTCTCTGCGCC 2100
QY 2101 CCTCTCAGCCAGGCGCATCAAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCT 2160
DB 2101 CCTCTCAGCCAGGCGCATCAAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCT 2160
QY 2161 GCAGCGAGCTGGGAGGAGGTTGAAACCGGCACTTACAGCAAGTGGGAGTCTGAAGAGC 2220
DB 2161 GCAGCGAGCTGGGAGGAGGTTGAAACCGGCACTTACAGCAAGTGGGAGTCTGAAGAGC 2220
QY 2221 CCTTGGAGGAGAAATATAAGAAACCAAGACATCCACCGCAATTAAGAGCAATTAACAC 2280
DB 2221 CCTTGGAGGAGAAATATAAGAAACCAAGACATCCACCGCAATTAAGAGCAATTAACAC 2280
QY 2281 CAGACCTTCCATGCCAGCGAGAGCTTTCCGCAAGGCGCCAGGCGCCCGGCGAGCT 2340
DB 2281 CAGACCTTCCATGCCAGCGAGAGCTTTCCGCAAGGCGCCAGGCGCCCGGCGAGCT 2340
QY 2341 GAGGAGCAACAGGCGAGAGCCCTTAAGCTCCAGGCTCTCTCCCAACAGAGCCCCCAGAG 2400
DB 2341 GAGGAGCAACAGGCGAGAGCCCTTAAGCTCCAGGCTCTCTCTCCCAACAGAGCCCCCAGAG 2400
QY 2401 CCAAAAGTCTCTCTCTTGA CTCTTGAAGAGGAGTCTGGGATGTGGGAAACCTTTA 2460

Db 2401 CCAAAACAAGCTCTCTCCCTTGAATTGTAGCAAGAGAGAGTCTGGGATGTGGGAAACCTTTA 2460
Qy 2461 CCTCTGTCTCTCCCTGAGAGCCAGCCCTGCGAGAAACCCAGCTCACAGAGCGGAAAGCA 2520
Db 2461 CCTCTGTCTCTCCCTGAGAGCCAGCCCTGCGAGAAACCCAGCTCACAGAGCGGAAAGCA 2520
Qy 2521 ACCGTCCGGAGCAGGAACTGAGAGAGTGGAAATAGAAATTATTCCTCAACAGACCTGTCC 2580
Db 2521 ACCGTCCGGAGCAGGAACTGAGAGAGTGGAAATAGAAATTATTCCTCAACAGACCTGTCC 2580
Qy 2581 CAGCCATTTTCTCTGGAGAGCAGGAAATTTCTCTGCTCAGATCGACAGCCTC 2640
Db 2581 CAGCCATTTTCTCTGGAGAGCAGGAAATTTCTCTGCTCAGATCGACAGCCTC 2640
Qy 2641 TCCCTGTCCGATGACAGTGGAGAACCCATCAAAAGGCTCTCAAGCTCCGGGACACC 2700
Db 2641 TCCCTGTCCGATGACAGTGGAGAACCCATCAAAAGGCTCTCAAAAGCTCCGGGACACC 2700
Qy 2701 CTGAGCTCAGGCGTACACTCTCTGGAGAGCCAGAGCGAGGCTCGAAGCTCCAGCTGGAAC 2760
Db 2701 CTGAGCTCAGGCGTACACTCTCTGGAGAGCCAGAGCGAGGCTCGAAGCTCCAGCTGGAAC 2760
Qy 2761 ATGCTGTGCGCCCGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAGTC 2820
Db 2761 ATGCTGTGCGCCCGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAGTC 2820
Qy 2821 CAAATACAGTCTCTTAATGGTGAAACCTGTGCACATCCGGAGTTCACCGGGTCAAGAGTG 2880
Db 2821 CAAATACAGTCTCTTAATGGTGAAACCTGTGCACATCCGGAGTTCACCGGGTCAAGAGTG 2880
Qy 2881 GGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCGAGCTTCAGCTTGTGTCAAC 2940
Db 2881 GGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCGAGCTTCAGCTTGTGTCAAC 2940
Qy 2941 AAAGAAGGAGCAGCTTTCGTACGACATGAGAGTGCCAGATCTCGGGCATCGACTGCGAG 3000
Db 2941 AAAGAAGGAGCAGCTTTCGTACGACATGAGAGTGCCAGATCTCGGGCATCGACTGCGAG 3000
Qy 3001 TGCACACTGGCCCTGATGGAGCTTCGCTGGAGCTGGAGGTCAGATGGCCAGCTG 3060
Db 3001 TGCACACTGGCCCTGATGGAGCTTCGCTGGAGCTGGAGGTCAGATGGCCAGCTG 3060
Qy 3061 GAGAACAGGCCCTTAACCTTGCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCTTC 3120
Db 3061 GAGAACAGGCCCTTAACCTTGCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCTTC 3120
Qy 3121 CTGCTCGGTGCAAGTGTGCTGCTGAAACACAGGCTCAGCGTTCAGGGGATGCGCA 3180
Db 3121 CTGCTCGGTGCAAGTGTGCTGCTGAAACACAGGCTCAGCGTTCAGGGGATGCGCA 3180
Qy 3181 GCGCCCGGCTCACAGTGGGAACAGGCGCTCGCAGCAGAGTGGGGCAAGCAGAT 3240
Db 3181 GCGCCCGGCTCACAGTGGGAACAGGCGCTCGCAGCAGAGTGGGGCAAGCAGAT 3240
Qy 3241 GCTTCCAGGATTTACACTCAGCCCTGCCCCACCTGCTGAAAGAAACATCCGCCAGT 3300
Db 3241 GCTTCCAGGATTTACACTCAGCCCTGCCCCACCTGCTGAAAGAAACATCCGCCAGT 3300
Qy 3301 GAAGAGACAGAGGAGTGGCAGGATTAACCTGGGGAAACAAACAGGGATCTTTTCT 3360
Db 3301 GAAGAGACAGAGGAGTGGCAGGATTAACCTGGGGAAACAAACAGGGATCTTTTCT 3360
Qy 3361 GCGCTGTCTCAGTCCAGTGGGCTGACCGCTTGGATCAGTGACCATTTGTGGCAGAC 3420
Db 3361 GCGCTGTCTCAGTCCAGTGGGCTGACCGCTTGGATCAGTGACCATTTGTGGCAGAC 3420
Qy 3421 AGGGGAGAGCAGCTTCAGCCTGGGTGAGAGGGTGGGCGAGCCCTTCGGCCCTCACC 3480
Db 3421 AGGGGAGAGCAGCTTCAGCCTGGGTGAGAGGGTGGGCGAGCCCTTCGGCCCTCACC 3480
Qy 3481 CTCAGGCTGTGTGAGAGTGTCAAGTGTGAAGGCCCAACCTCAGGTTTCAGTGAGAA 3540
Db 3481 CTCAGGCTGTGTGAGAGTGTCAAGTGTGAAGGCCCAACCTCAGGTTTCAGTGAGAA 3540

Qy 3541 CCAGTCAGCAGGTATGCCCGCTTAGGTTAAGGGGCCCTCTAAACCCCTTGCTGGC 3600
Db 3541 CCAGTCAGCAGGTATGCCCGCTTAGGTTAAGGGGCCCTCTAAACCCCTTGCTGGC 3600
Qy 3601 CTCACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAGATGCTGAGACCTGGGAAG 3660
Db 3601 CTCACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAGATGCTGAGACCTGGGAAG 3660
Qy 3661 GCTCCCTGGTGAATACACACACTTTTCAGGTTGTGCAACACAGGTCCTGAGTTGACC 3720
Db 3661 GCTCCCTGGTGAATACACACACTTTTCAGGTTGTGCAACACAGGTCCTGAGTTGACC 3720
Qy 3721 TCTGTTTCAGCAAGGACCAAGAGGTGTGTAAAGTGAAGTGTCTCAGTCCCCCAGACA 3780
Db 3721 TCTGTTTCAGCAAGGACCAAGAGGTGTGTAAAGTGAAGTGTCTCAGTCCCCCAGACA 3780
Qy 3781 TGTGCCCTTTTGTGCTGCTTACCACTCTTCCCAGAGCAGAGGCCCGAGCCCTTCA 3840
Db 3781 TGTGCCCTTTTGTGCTGCTTACCACTCTTCCCAGAGCAGAGGCCCGAGCCCTTCA 3840
Qy 3841 GGGCCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGT 3900
Db 3841 GGGCCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGT 3900
Qy 3901 GATGAGGATATGCTGACAGAAACAGTCTCTGGATGGACATGATCAGTGTAAAGAAAG 3960
Db 3901 GATGAGGATATGCTGACAGAAACAGTCTCTGGATGGACATGATCAGTGTAAAGAAAG 3960
Qy 3961 CAGCAGAGAGAGAGTCCGGCGCCAGCCCACTATCAGTGTCCAGCGTGTGTTCC 4020
Db 3961 CAGCAGAGAGAGAGTCCGGCGCCAGCCCACTATCAGTGTCCAGCGTGTGTTCC 4020
Qy 4021 CAGAGCAAGCTCAGCATCACACTGACACTCACCTGCGCTGCCCTGGCCAGAGGGTAC 4080
Db 4021 CAGAGCAAGCTCAGCATCACACTGACACTCACCTGCGCTGCCCTGGCCAGAGGGTAC 4080
Qy 4081 TGCCGACGGCACTTTTGCACTCTGATGACCTCAAAAGCACTTTTCATGGCTGCCCTTGGCAG 4140
Db 4081 TGCCGACGGCACTTTTGCACTCTGATGACCTCAAAAGCACTTTTCATGGCTGCCCTTGGCAG 4140
Qy 4141 GGCAGGCGAGGCGAGTGACACTGTAGGAGCATAGCAAGCAGAGAGATGGGTGAAGGAC 4200
Db 4141 GGCAGGCGAGGCGAGTGACACTGTAGGAGCATAGCAAGCAGAGAGATGGGTGAAGGAC 4200
Qy 4201 ACAGTCTTTGAGCTGTCCACATGCTGACTCCTCAAACTCTCCAGATTCTCTAAGA 4260
Db 4201 ACAGTCTTTGAGCTGTCCACATGCTGACTCCTCAAACTCTCCAGATTCTCTAAGA 4260
Qy 4261 ATAGCACCCCTTTCCCATTTGCCAGCTTAGCCTCTTCTCCAGGGGAGCTACTCAGGA 4320
Db 4261 ATAGCACCCCTTTCCCATTTGCCAGCTTAGCCTCTTCTCCAGGGGAGCTACTCAGGA 4320
Qy 4321 CTCAGTAGCATTAATCAGCTGTGAATCGTTCAGGGGTGTCTGCTAGCTCAACCTCT 4380
Db 4321 CTCAGTAGCATTAATCAGCTGTGAATCGTTCAGGGGTGTCTGCTAGCTCAACCTCT 4380
Qy 4381 GGGGCGAGGGGAGCGGAGACTCCGTGGGAGAGCTCATTTCCACATCTTGCCAGACAGC 4440
Db 4381 GGGGCGAGGGGAGCGGAGACTCCGTGGGAGAGCTCATTTCCACATCTTGCCAGACAGC 4440
Qy 4441 CTTTCTCCAGCTGTCCACATTAAGTCAGTCTGCTCCCGGGGAGAGGCCCGGCCCCAG 4500
Db 4441 CTTTCTCCAGCTGTCCACATTAAGTCAGTCTGCTCCCGGGGAGAGGCCCGGCCCCAG 4500
Qy 4501 CACATAAAGAACTGAGCCTTGGTACTGAGAGTCTGGGTGTGAGAGAACTTTTGTAAAG 4560
Db 4501 CACATAAAGAACTGAGCCTTGGTACTGAGAGTCTGGGTGTGAGAGAACTTTTGTAAAG 4560
Qy 4561 CAATAAAGTTTGGGTGATGACAAATGTTTAAAAA 4596
Db 4561 CAATAAAGTTTGGGTGATGACAAATGTTTAAAAA 4596

RESULT 2

US-09-155-676A-6
; Sequence 6, Application US/09155676A
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
; TITLE OF INVENTION: FACTOR (TRAP), THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/155,676A
; FILING DATE: 04-JAN-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IL97/00117
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117800
; FILING DATE: 02-APR-1996
; APPLICATION NUMBER: IL 119133
; FILING DATE: 26-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: WALLACH=21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4596 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-155-676A-6

Query Match - 100.0%; Score 4596; DB 17; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGCGGGGGGACTGTGCGGTGGAAACGTGTAGCTGTTGAAGGTGGACTCTGTTACCATTTG 60
Db 1 AGCGGGGGGACTGTGCGGTGGAAACGTGTAGCTGTTGAAGGTGGACTCTGTTACCATTTG 60
Qy 61 AGGATGTTTGAGGATGAGTATGTGTGCGCAGGCAACATATAACAGGACAGACCCCTTT 120
Db 61 AGGATGTTTGAGGATGAGTATGTGTGCGCAGGCAACATATAACAGGACAGACCCCTTT 120
Qy 121 GCGCTGCTTTTCCCAACCAAGCTGACCTGTGTTCTCCAGGTCTGGGATTTCTA 180
Db 121 GCGCTGCTTTTCCCAACCAAGCTGACCTGTGTTCTCCAGGTCTGGGATTTCTA 180
Qy 181 AGTGACCTGCTCTGTGTTTGTCTCTCTCAGGATGAGCACAAAGCTGGGAGATGGCAGTG 240
Db 181 AGTGACCTGCTCTGTGTTTGTCTCTCTCAGGATGAGCACAAAGCTGGGAGATGGCAGTG 240

Qy 241 ATGGAATGCGCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGAGAACTCCCC 300
Db 241 ATGGAATGCGCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGAGAACTCCCC 300
Qy 301 AAGCCAAAGGAGAGACGCGCCACTGGGGAAGAAACAGAGCTCGTCTACAAGTTGAG 360
Db 301 AAGCCAAAGGAGAGACGCGCCACTGGGGAAGAAACAGAGCTCGTCTACAAGTTGAG 360
Qy 361 GCGTGGAGAAAGAGCCCTGTGTTCTGCGGAAAGTGGAGATCCTGAATGAGTGAATACC 420
Db 361 GCGTGGAGAAAGAGCCCTGTGTTCTGCGGAAAGTGGAGATCCTGAATGAGTGAATACC 420
Qy 421 AAGGGCAGCAGCAAGGAAGGCTCCGAGGCGAGGGCAGCTGCCATCTCTATCATCGCCAG 480
Db 421 AAGGGCAGCAGCAAGGAAGGCTCCGAGGCGAGGGCAGCTGCCATCTCTATCATCGCCAG 480
Qy 481 GCTGAGTGTGAGAAATAGCCAAAGAGTTCAGCCCCACCTTTTTCAGAAAGGATTTTATCGCT 540
Db 481 GCTGAGTGTGAGAAATAGCCAAAGAGTTCAGCCCCACCTTTTTCAGAAAGGATTTTATCGCT 540
Qy 541 GGGTCCAAACAGTACAGCAGTCCGAGAGTCTTCATCAGATCCCCCAACAATGTGSCCCAT 600
Db 541 GGGTCCAAACAGTACAGCAGTCCGAGAGTCTTCATCAGATCCCCCAACAATGTGSCCCAT 600
Qy 601 GCTACAGAGGGCAAAATGGGCCGTGTGTTGGGAAGGAAGCGTCGAGCAAAAGCCCGG 660
Db 601 GCTACAGAGGGCAAAATGGGCCGTGTGTTGGGAAGGAAGCGTCGAGCAAAAGCCCGG 660
Qy 661 AAGAAAACGGAAGAAAGAGCTCAAAAGTCCCTGCTCATGAGGAGTGGCTTTGGCATA 720
Db 661 AAGAAAACGGAAGAAAGAGCTCAAAAGTCCCTGCTCATGAGGAGTGGCTTTGGCATA 720
Qy 721 CCCCTCCAGAGCCCTGAGCAGGAGCTGACCATCCCAAGTGGAGGAGTGAAGTCT 780
Db 721 CCCCTCCAGAGCCCTGAGCAGGAGCTGACCATCCCAAGTGGAGGAGTGAAGTCT 780
Qy 781 CCAGTCCGCGCCCATATGTTAGAAAACCCCGCAGTTCACCAAGCTCTCAAGGAACA 840
Db 781 CCAGTCCGCGCCCATATGTTAGAAAACCCCGCAGTTCACCAAGCTCTCAAGGAACA 840
Qy 841 GGCCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGGCGGCTCTGCTCGA 900
Db 841 GGCCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGGCGGCTCTGCTCGA 900
Qy 901 TCAGAACTCCACAACTGATCAGCCCTTGAATGTTGAACCAAGTGTGAAACTGAC 960
Db 901 TCAGAACTCCACAACTGATCAGCCCTTGAATGTTGAACCAAGTGTGAAACTGAC 960
Qy 961 CACCCAGGAGCGGAGCGCCCTGCGCCCTGCGCCAGCAGCCCTTCCCTATAGCAGACTG 1020
Db 961 CACCCAGGAGCGGAGCGCCCTGCGCCCTGCGCCAGCAGCCCTTCCCTATAGCAGACTG 1020
Qy 1021 CCTCATCCCTTCCCATTCACCTCTCCAGCCCTGGAAACCTCACCTCTGGAGTCTTTC 1080
Db 1021 CCTCATCCCTTCCCATTCACCTCTCCAGCCCTGGAAACCTCACCTCTGGAGTCTTTC 1080
Qy 1081 CTGGGCAAACTGGGCTGTGTAGACAGCAGAAACCTTGTGCTGACCCACACTGAGCAAA 1140
Db 1081 CTGGGCAAACTGGGCTGTGTAGACAGCAGAAACCTTGTGCTGACCCACACTGAGCAAA 1140
Qy 1141 CTGGCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTTGGAGCCAGCTGCTG 1200
Db 1141 CTGGCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTTGGAGCCAGCTGCTG 1200
Qy 1201 TCTGCTGGTGGCCATGAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGC 1260
Db 1201 TCTGCTGGTGGCCATGAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGC 1260
Qy 1261 AGCGTGAGCTCAAGCCAGGCCCCACAGCTGACAGCTGGCCAAAGACTGGGAGCAGCG 1320
Db 1261 AGCGTGAGCTCAAGCCAGGCCCCACAGCTGACAGCTGGCCAAAGACTGGGAGCAGCG 1320

1321 GGTCTCCAGATCCCGGAGCCCAAGCCCCAAAATGAGGACAAAGAGGGTGTCTGCTCACT 1380
1321 GGTCTCCAGATCCCGGAGCCCAAGCCCCAAAATGAGGACAAAGAGGGTGTCTGCTCACT 1380
1381 GAGAACTCAAGCCAGTGGATTATGAGTACCGAGAAGAAAGTCCACTGGGCCACGACCCAG 1440
1381 GAGAACTCAAGCCAGTGGATTATGAGTACCGAGAAGAAAGTCCACTGGGCCACGACCCAG 1440
1441 CTCGCGCTGGGAGAGGGCTTCCTTCGGAGAGGTGACAGGATGGAGGACAAAGCAGACTGGC 1500
1441 CTCGCGCTGGGAGAGGGCTTCCTTCGGAGAGGTGACAGGATGGAGGACAAAGCAGACTGGC 1500
1501 TTCAGTGGCTGTCAAAAGGTGGCCCTGGGAAGTATTTTCGGGCAGAGGAGTGTATGGCA 1560
1501 TTCAGTGGCTGTCAAAAGGTGGCCCTGGGAAGTATTTTCGGGCAGAGGAGTGTATGGCA 1560
1561 TGTGAGGATTCACCTCACCAGAAATTTGTCCTTTGTATGGAGCTGTGAGAGAGGGCTT 1620
1561 TGTGAGGATTCACCTCACCAGAAATTTGTCCTTTGTATGGAGCTGTGAGAGAGGGCTT 1620
1621 TGGGTCAACATCTTCATGAGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTGCTCAAGGAG 1680
1621 TGGGTCAACATCTTCATGAGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTGCTCAAGGAG 1680
1681 CAGGGCTGTCTCCAGAGACCGGGCCCTGTACTACCTGGGCCAGGCCCTCGAGGGTCTG 1740
1681 CAGGGCTGTCTCCAGAGACCGGGCCCTGTACTACCTGGGCCAGGCCCTCGAGGGTCTG 1740
1741 GAATACCTTCACACGAAAGGATTCGCATGGGGACGTCGCAAGCTGACAAAGCTGACAA 1800
1741 GAATACCTTCACACGAAAGGATTCGCATGGGGACGTCGCAAGCTGACAAAGCTGACAA 1800
1801 TCCAGGATGGAGACGACGAGCCCTCTGTGACTTTGGCCATGCTGTGCTTCAAGCT 1860
1801 TCCAGGATGGAGACGACGAGCCCTCTGTGACTTTGGCCATGCTGTGCTTCAAGCT 1860
1861 GATGGCTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGGCACAGAGACCCACATG 1920
1861 GATGGCTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGGCACAGAGACCCACATG 1920
1921 GCTCGGAGGTGGTGTGGGACGAGTGCAGCGCCAAAGGTGGATGTCTGGAGCAGCTGC 1980
1921 GCTCGGAGGTGGTGTGGGACGAGTGCAGCGCCAAAGGTGGATGTCTGGAGCAGCTGC 1980
1981 TGTATGATGCTGCATGCTCAACGCTGACCGCTGACCTGACCTGACCTGACCTGACCTG 2040
1981 TGTATGATGCTGCATGCTCAACGCTGACCGCTGACCTGACCTGACCTGACCTGACCTG 2040
2041 CTCCTGCTCAAGATTGCCAGCGAGCTCCGCTGTGAGGGAGATCCACCTCTCTGGCC 2100
2041 CTCCTGCTCAAGATTGCCAGCGAGCTCCGCTGTGAGGGAGATCCACCTCTCTGGCC 2100
2101 CCTCTCAGCCAGGCCATCCAGAGGGGCTGAGGAAAGGCCCATCCACCGCTGTCT 2160
2101 CCTCTCAGCCAGGCCATCCAGAGGGGCTGAGGAAAGGCCCATCCACCGCTGTCT 2160
2161 GCAGCGGAGCTGGGAGGAGTGAACCGGGCACTACAGCAAGTGGAGGCTCTGAAGAGC 2220
2161 GCAGCGGAGCTGGGAGGAGTGAACCGGGCACTACAGCAAGTGGAGGCTCTGAAGAGC 2220
2221 CTTTGGAGGGAGATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAC 2280
2221 CTTTGGAGGGAGATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAC 2280
2281 CAGACCTTCATGCCAGCCGAGAGCTTTTCGCAAGGGCCCCAGGGCCCCCGGCGAGCT 2340
2281 CAGACCTTCATGCCAGCCGAGAGCTTTTCGCAAGGGCCCCAGGGCCCCCGGCGAGCT 2340
2341 GAGGAGACAAAGGAGGAGCCCTTACGCTCCAGCTCCTCTCCACCCAGAGCCCCCAGAG 2400
2341 GAGGAGACAAAGGAGGAGCCCTTACGCTCCAGCTCCTCTCTCCACCCAGAGCCCCCAGAG 2400
2401 CCAAAACAAGTCTCTCTCTTGTAGCTTTTGTAGCAAGGAGGAGTCTGGGATGTGGGAAACCTTTA 2460

2401 CCAAAACAAGTCTCTCTCTTGTAGCTTTGTAGCAAGGAGGAGTCTGGGATGTGGGAAACCTTTA 2460
2461 CCTGTGTCTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACCAGAGCGGAAAGCA 2520
2461 CCTGTGTCTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACCAGAGCGGAAAGCA 2520
2521 ACCGTCCCGGAGCAGGAACTGAGCAGCTGCAGAAATAGAAATATTCTCTCAACAGCCTGTCC 2580
2521 ACCGTCCCGGAGCAGGAACTGAGCAGCTGCAGAAATAGAAATATTCTCTCAACAGCCTGTCC 2580
2581 CAGCCATTTCTCTGGAGGACAGGACAAATTTCTCTCGTGCCTCAGCATCGACAGCCTC 2640
2581 CAGCCATTTCTCTGGAGGACAGGACAAATTTCTCTCGTGCCTCAGCATCGACAGCCTC 2640
2641 TCCCTGTCCGATGACAGTGAGNAACCCATCAAGAGCCTCTCAAGCTCGCGGACACC 2700
2641 TCCCTGTCCGATGACAGTGAGNAACCCATCAAGAGCCTCTCAAGCTCGCGGACACC 2700
2701 CTGAGCTCAGGCTACACTCTCTGGAGCAGCAGGCGGAGGCTCGAAGCTCCAGCTGGAAAC 2760
2701 CTGAGCTCAGGCTACACTCTCTGGAGCAGCAGGCGGAGGCTCGAAGCTCCAGCTGGAAAC 2760
2761 ATGCTGTGGCCCGGGGCGGCCACCAACCCCAAGCTATTTCAATGGTGTGAAAGTCC 2820
2761 ATGCTGTGGCCCGGGGCGGCCACCAACCCCAAGCTATTTCAATGGTGTGAAAGTCC 2820
2821 CAATACAGTCTCTTAATGGTGTGACACCTGACATCCGGGAGTTCACCGGGTCAAGTCC 2880
2821 CAATACAGTCTCTTAATGGTGTGACACCTGACATCCGGGAGTTCACCGGGTCAAGTCC 2880
2881 GGAGACATCGGCACCTGAGCATCAGCAGCAGATCCAGCTGAGCCTTCAGCTTGTGTACCC 2940
2881 GGAGACATCGGCACCTGAGCATCAGCAGCAGATCCAGCTGAGCCTTCAGCTTGTGTACCC 2940
2941 AAAGACGGGACGCTTTCGCTACGATGAGGAGTGCAGACTCGGGCATCGACCTGTGCTCAG 3000
2941 AAAGACGGGACGCTTTCGCTACGATGAGGAGTGCAGACTCGGGCATCGACCTGTGCTCAG 3000
3001 TGCACATGGGCCCTGATGGCAGCTTCGCTGAGCTGGAGGTCAGCATGGCCAGCTG 3060
3001 TGCACATGGGCCCTGATGGCAGCTTCGCTGAGCTGGAGGTCAGCATGGCCAGCTG 3060
3061 GAGAACGGCCCTAACCTGCGCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCTTC 3120
3061 GAGAACGGCCCTAACCTGCGCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCTTC 3120
3121 CTGCTCGGTGACAGATGTGCTGTAACACAGGCTCAGCGCTTCAGGGGATTTGCCA 3180
3121 CTGCTCGGTGACAGATGTGCTGTAACACAGGCTCAGCGCTTCAGGGGATTTGCCA 3180
3181 GCGCCCGGCTCAGAGTGGGAAACAGGGCTCGAGCAGCAGAGTGGGGGCAAGCAGAT 3240
3181 GCGCCCGGCTCAGAGTGGGAAACAGGGCTCGAGCAGCAGAGTGGGGGCAAGCAGAT 3240
3241 GCGTCCAGGATTTTCAACCTGAGCCTGCGCCACCTCTGCTGAAACAAATCCCGCAGCT 3300
3241 GCGTCCAGGATTTTCAACCTGAGCCTGCGCCACCTCTGCTGAAACAAATCCCGCAGCT 3300
3301 GAAGAGACAGAGGAGGATGGCAGGATTAAGCTGGGGAAACAAACAGGGATCTTTTCT 3360
3301 GAAGAGACAGAGGAGGATGGCAGGATTAAGCTGGGGAAACAAACAGGGATCTTTTCT 3360
3361 GCGCCCTGCTCCAGTCCAGTTCGCTGACCGCTGGATTCAGTTCAGTTCAGTTCAGTTCAG 3420
3361 GCGCCCTGCTCCAGTCCAGTTCGCTGACCGCTGGATTCAGTTCAGTTCAGTTCAGTTCAG 3420
3421 AGGGAGAGCAGCTTCAGCCTGGGTCAAGAGGGGTGGGCGAGCCTTCGGGCCCTCACC 3480
3421 AGGGAGAGCAGCTTCAGCCTGGGTCAAGAGGGGTGGGCGAGCCTTCGGGCCCTCACC 3480
3481 CTCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGGCCCAAACTCAGGTTCTAGTGCAGAA 3540

Db 3481 CTCCAGGCTGCTGAGAGTGTCAAGTGTGTAAGGGCCCAAACTCAGGTTCAAGTCAGAA 3540
QY 3541 CCAGGTACAGAGTATGCCCGCCCGTAGGTTAAGGGGCCCTCTAAACCCCTTGCTGGC 3600
Db 3541 CCAGGTACAGAGTATGCCCGCCCGTAGGTTAAGGGGCCCTCTAAACCCCTTGCTGGC 3600
QY 3601 CTCACCTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAGAAATCCCTGACCTGGGAAG 3660
Db 3601 CTCACCTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAGAAATCCCTGACCTGGGAAG 3660
QY 3661 GCTCCCTGGTAGAATACACCACTTTTTCAGGTTGTGCAACACACAGGTCCTGAGTTGACC 3720
Db 3661 GCTCCCTGGTAGAATACACCACTTTTTCAGGTTGTGCAACACACAGGTCCTGAGTTGACC 3720
QY 3721 TCTGTTTCAGCCAGGACCAAGAGAGGTGTGTAAGTGAAGTGGTTCTCAGTCCCCAGACA 3780
Db 3721 TCTGTTTCAGCCAGGACCAAGAGAGGTGTGTAAGTGAAGTGGTTCTCAGTCCCCAGACA 3780
QY 3781 TGTGCCCTTTTGTCTGGCTACCACTCTTCCCAAGAGCAGGCGCCGAGCCCTTCA 3840
Db 3781 TGTGCCCTTTTGTCTGGCTACCACTCTTCCCAAGAGCAGGCGCCGAGCCCTTCA 3840
QY 3841 GCGCCAGCACTGCCCGCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGTGAAGGT 3900
Db 3841 GCGCCAGCACTGCCCGCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGTGAAGGT 3900
QY 3901 GATGAGGATATGCTCTGACAGAAACAGTCTGTGGATGACATGATCAGTGTCAAGGAAAG 3960
Db 3901 GATGAGGATATGCTCTGACAGAAACAGTCTGTGGATGACATGATCAGTGTCAAGGAAAG 3960
QY 3961 CAGCAGAGAGAGCTCGCGGCCCGCCAGCCCACTATCAGTGTCCAGGTCGTGGTTCCC 4020
Db 3961 CAGCAGAGAGAGAGCTCGCGGCCCGCCAGCCCACTATCAGTGTCCAGGTCGTGGTTCCC 4020
QY 4021 CAGAGCAGCTCAGCATCACACTGACACTCACCTGCCCTGCCCTGCCCGCAGAGGATC 4080
Db 4021 CAGAGCAGCTCAGCATCACACTGACACTCACCTGCCCTGCCCTGCCCGCAGAGGATC 4080
QY 4081 TGCCGACGGCACTTTGGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTGGCAG 4140
Db 4081 TGCCGACGGCACTTTGGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTGGCAG 4140
QY 4141 GGCAGGCGAGGCGAGTACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGAC 4200
Db 4141 GGCAGGCGAGGCGAGTACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGAC 4200
QY 4201 ACAGTCTTTGAGCTGTCCACATGATGCTGACTCTCCCAACCTCTTCCAGATTTCTTAAGA 4260
Db 4201 ACAGTCTTTGAGCTGTCCACATGATGCTGACTCTCCCAACCTCTTCCAGATTTCTTAAGA 4260
QY 4261 ATAGCACCCTTTCCCAATTCGCCAGCTTAGCCTCTTCTCCAGGGGAGACTACTCAGGA 4320
Db 4261 ATAGCACCCTTTCCCAATTCGCCAGCTTAGCCTCTTCTCCAGGGGAGACTACTCAGGA 4320
QY 4321 CTCAGTAGCATTAATCAGCTGTGAATCGTCAGGGGTGTCTGTAGCCTCAACCTCCT 4380
Db 4321 CTCAGTAGCATTAATCAGCTGTGAATCGTCAGGGGTGTCTGTAGCCTCAACCTCCT 4380
QY 4381 GGGGAGGGGAGCGGAGACTCCGTGGGAGAGCTCATTTCCCACTCTTGGCCAGAGACGC 4440
Db 4381 GGGGAGGGGAGCGGAGACTCCGTGGGAGAGCTCATTTCCCACTCTTGGCCAGAGACGC 4440
QY 4441 CTTTGTCCAGCTGTCCCAATTCGCCAGCTTAGCCTCTTCTCCAGGGGAGAGCGCCGGCCCCCAG 4500
Db 4441 CTTTGTCCAGCTGTCCCAATTCGCCAGCTTAGCCTCTTCTCCAGGGGAGAGCGCCGGCCCCCAG 4500
QY 4501 CACATAAGAACTGCAGCTTTGGTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTTGTAAAG 4560
Db 4501 CACATAAGAACTGCAGCTTTGGTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTTGTAAAG 4560
QY 4561 CAATAAAGTTTGGGTGTAGTCAATGTTTAAAAAA 4596
Db 4561 CAATAAAGTTTGGGTGTAGTCAATGTTTAAAAAA 4596

RESULT 3

PCT-US01-08631-15294
; Sequence 15294, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 15294
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1117)..(1581)
; OTHER INFORMATION: 100% homologous to Homo sapiens sperm protein 10, accession
; OTHER INFORMATION: number M82968, Smith-Waterman Score=809.
PCT-US01-08631-15294

Query Match 99.6%; Score 4578.4; DB 1; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 5 GGGGAGCTGTGCCGTGTGGAACTGTAGCTCTTTGA-AGGTGGACTCTCTTACCATTTGAGG 63
Db 5 GGGGAGCTGTGCCGTGTGGAACTGTAGCTCTTTGAGAGTGGACTCTCTTACCATTTGAGG 64
QY 64 ATGTTTGAGGATGAGTATGTGTGGCAGAGCAACATAAAACAGCAGAGACCTTTTGGC 123
Db 65 ATGTTTGAGGATGAGTATGTGTGGCAGAGCAACATAAAACAGCAGAGACCTTTTGGC 124
QY 124 CCTGCCCTTCTCCCGCAACCCAGGCTGACCTGTGTTCTCCAGGCTCTGGGATTTCTAAGT 183
Db 125 CCTGCCCTTCTCCCGCAACCCAGGCTGACCTGTGTTCTCCAGGCTCTGGGATTTCTAAGT 184
QY 184 GACCTGCTCTGTGTTTGTGTTCTCTCAGGATGAGCACAAGCTTGGAGATGGCAGTGTATG 243
Db 185 GACCTGCTCTGTGTTTGTGTTCTCTCAGGATGAGCACAAGCTTGGAGATGGCAGTGTATG 244
QY 244 GAAATGGCCTGCCCGCAGGTCGCCCTTGGCTCAGCAGTGGGGCAGCAGAGAACTCCCCAAG 303
Db 245 GAAATGGCCTGCCCGCAGGTCGCCCTTGGCTCAGCAGTGGGGCAGCAGAGAACTCCCCAAG 304
QY 304 CCAAGGAGAGAGCGCCGCACTGGGAGAGAAACAGAGCTCGGTCTACAAGCTTGAGGCC 363
Db 305 CCAAGGAGAGAGCGCCGCACTGGGAGAGAAACAGAGCTCGGTCTACAAGCTTGAGGCC 364
QY 364 GTGAGAGAGAGCGCTGTGTTCTGCGGAAAGTGGGAGATCCCTGAATGAGCTGATTACCAAG 423
Db 365 GTGAGAGAGAGCGCTGTGTTCTGCGGAAAGTGGGAGATCCCTGAATGAGCTGATTACCAAG 424
QY 424 GGCAACAGCAAGAGGCTCCGAGGCGAGGCGAGCTGCCATCTCTATCATCGCCAGGCT 483
Db 425 GGCAACAGCAAGAGGCTCCGAGGCGAGGCGAGCTGCCATCTCTATCATCGCCAGGCT 484
QY 484 GAGTGTAGATAGCCAGAGTTTTCAGCCCACTTTTTCAGAACCGATTTTTCATCGCTGGG 543
Db 485 GAGTGTAGATAGCCAGAGTTTTCAGCCCACTTTTTCAGAACCGATTTTTCATCGCTGGG 544
QY 544 TCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCCATGCT 603
Db 545 TCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCCATGCT 604
QY 604 ACAGAGGGCAAAATGGCCCGTGTGTGTTGGAGGAAAGCGCTCGCAGCAAAAGCCCGGAAG 663

605	Db	 ACAGAGGGCAAAATGGCCCGTGTGTGTGTAAGAGGGAAGCGTCCGACAGAAAGCCCGGAAG	664
664	Qy	 AAACGGAAAGAAAGAGCTCAAAGTCCCTGGCTCATGACGAGTGGCCCTTGGGCCAAACCC	723
665	Db	 AAACGGAAAGAAAGAGCTCAAAGTCCCTGGCTCATGACGAGTGGCCCTTGGGCCAAACCC	724
724	Qy	 CTCCCCAGGACCCCTGTAGCAGGAGAGCTGCACCATCCCAAGCCTCTGAAGGAACAGGC	783
725	Db	 CTCCCCAGGACCCCTGTAGCAGGAGAGCTGCACCATCCCAAGCCTCTGAAGGAACAGGC	784
784	Qy	 CTGGGCGCCCATATATGTTAGAAACACCCCGAGTTCAACAAAGCCTCTGAAGGAACAGGC	843
785	Db	 CTGGGCGCCCATATATGTTAGAAACACCCCGAGTTCAACAAAGCCTCTGAAGGAACAGGC	844
844	Qy	 CTTGGGCAACTCTGTTTTAAGCAGCTTGGCGAGGGCCTACGGCGGCTCTGCCTCGCATCA	903
845	Db	 CTTGGGCAACTCTGTTTTAAGCAGCTTGGCGAGGGCCTACGGCGGCTCTGCCTCGCATCA	904
904	Qy	 GAACTCCACAAACTGATCAGCCCCCTTGCAATGTCTGAACCAACGCTGTGGAAATGCAACC	963
905	Db	 GAACTCCACAAACTGATCAGCCCCCTTGCAATGTCTGAACCAACGCTGTGGAAATGCAACC	964
964	Qy	 CCCAGAACGGAGGCCCCCTGCCCCCTGCCACGACCCCTTCCCTCTATAGCAACTGCCT	1023
965	Db	 CCCAGAACGGAGGCCCCCTGCCCCCTGCCACGACCCCTTCCCTCTATAGCAACTGCCT	1024
1024	Qy	 CATCCCTTCCATTTCCACCCCTCTCCAGCCCTGGAAACCTCACCTCTGGAGTCTTCCCTG	1083
1025	Db	 CATCCCTTCCATTTCCACCCCTCTCCAGCCCTGGAAACCTCACCTCTGGAGTCTTCCCTG	1084
1084	Qy	 GGCAAACTGGCCTGTGTAGCAGCGCAAAACCCCTTGCTGACCCACACCTTGAGCAAACTG	1143
1085	Db	 GGCAAACTGGCCTGTGTAGCAGCGCAAAACCCCTTGCTGACCCACACCTTGAGCAAACTG	1144
1144	Qy	 GCCTGTGTAGACAGTCCAAAGCCCTGCCTGGGCCACACTTGGAGGCCAGCTGCCTGTCT	1203
1145	Db	 GCCTGTGTAGACAGTCCAAAGCCCTGCCTGGGCCACACTTGGAGGCCAGCTGCCTGTCT	1204
1204	Qy	 CGTGGTCCCATGAGAGATTTCTGTGGAGGATACCTAGTGCATGCTCTGCAAGGCAGC	1263
1205	Db	 CGTGGTCCCATGAGAGATTTCTGTGGAGGATACCTAGTGCATGCTCTGCAAGGCAGC	1264
1264	Qy	 GTGAGCTCAAGCCAGGCCACAGCCTGACAGCCTGGCCAGAGCCTGGGCAGACAGGGGC	1323
1265	Db	 GTGAGCTCAAGCCAGGCCACAGCCTGACAGCCTGGCCAGAGCCTGGGCAGACAGGGGC	1324
1324	Qy	 TCCAGATCCCGGGAGCCAGCCCAAACTGAGGACAAACGAGGCTGCTCTGCTCACTGAG	1383
1325	Db	 TCCAGATCCCGGGAGCCAGCCCAAACTGAGGACAAACGAGGCTGCTCTGCTCACTGAG	1384
1384	Qy	 AAACTCAAGCCAGTGGATATATGAGTACGAGAGAAGTCCACTGGGCCACGCACCAAGCTC	1443
1385	Db	 AAACTCAAGCCAGTGGATATATGAGTACGAGAGAAGTCCACTGGGCCACGCACCAAGCTC	1444
1444	Qy	 CGCCTGGGCAGAGGCTCCTTCGGAGAGGTGCACAGGATGAGGACAAGAGACTGGCTTC	1503
1445	Db	 CGCCTGGGCAGAGGCTCCTTCGGAGAGGTGCACAGGATGAGGACAAGAGACTGGCTTC	1504
1504	Qy	 CAGTGCCTGTCAAAAAGTTCGGAAGTATTTTCGGGCAGAGGAGCTGATGGCATGT	1563
1505	Db	 CAGTGCCTGTCAAAAAGTTCGGAAGTATTTTCGGGCAGAGGAGCTGATGGCATGT	1564
1564	Qy	 GCAGATTGACCTCACCAGAAATTTGCTTGTATGGAGCTGTGAGAGAAGGCCCTTGG	1623
1565	Db	 GCAGATTGACCTCACCAGAAATTTGCTTGTATGGAGCTGTGAGAGAAGGCCCTTGG	1624
1624	Qy	 GTCACATCTTCATGGAGCTCCTCGAAGGTGGCTCCCTGGGCCAGCTGGTCAAGAGACAG	1683
1625	Db	 GTCACATCTTCATGGAGCTCCTCGAAGGTGGCTCCCTGGGCCAGCTGGTCAAGAGACAG	1684
1684	Qy	 GGCTGTCTTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGCCCTTGGAGGCTCGAA	1743

1585	DB	GGCTGTCTCCAGAGGACGGGGCCCTGTACTACCTGGGCCAGGCCCTGGAGGGTCTGGAA	1744
1744	QY	TACCTCCACTCAGAAAGGATTTCTGATGGGGACGTCAAAAGCTGACAAAGTCTCTGTGTC	1803
1745	DB	TACCTCCACTCAGAAAGGATTTCTGATGGGGACGTCAAAAGCTGACAAAGTCTCTGTGTC	1804
1804	QY	AGCGATGGGAGCCACGAGGCCCTCTGTGACTTTGGCCATGCTGTGTCTTTCAACCTGAT	1863
1805	DB	AGCGATGGGAGCCACGAGGCCCTCTGTGACTTTGGCCATGCTGTGTCTTTCAACCTGAT	1864
1864	QY	GGCTCTGGAAAGTCTTGTCTACAGGGGACTACATCCCTGGCACAGAGACCCACATGGCT	1923
1865	DB	GGCTCTGGAAAGTCTTGTCTACAGGGGACTACATCCCTGGCACAGAGACCCACATGGCT	1924
1924	QY	CCGGAGTGTGTCTGGSCAGGAGCTGCGACGCCAAGGTGGATGCTCTGGAGCAGCTGCTGT	1983
1925	DB	CCGGAGTGTGTCTGGSCAGGAGCTGCGACGCCAAGGTGGATGCTCTGGAGCAGCTGCTGT	1984
1984	QY	ATGATGTCGACATGCTCAAACGGGTGCCACCCCTTGACTCAGTTCCTCGAGGGCCGCTC	2043
1985	DB	ATGATGTCGACATGCTCAAACGGGTGCCACCCCTTGACTCAGTTCCTCGAGGGCCGCTC	2044
2044	QY	TGCTCTCAAGATTGCCACGAGCCCTCCGCTGTGAGGGAGATCCACCCCTCTCGGCCCT	2103
2045	DB	TGCTCTCAAGATTGCCACGAGCCCTCCGCTGTGAGGGAGATCCACCCCTCTCGGCCCT	2104
2104	QY	CTCACAGCCAGGCCATCTCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGCA	2163
2105	DB	CTCACAGCCAGGCCATCTCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGCA	2164
2164	QY	GCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGGCCCT	2223
2165	DB	GCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGGCCCT	2224
2224	QY	TGGAGGGGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACCACCG	2283
2225	DB	TGGAGGGGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACCACCG	2284
2284	QY	ACCTCTCATGCCCCAGCCAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCCGCCAGCTGAG	2343
2285	DB	ACCTCTCATGCCCCAGCCAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCCGCCAGCTGAG	2344
2344	QY	GAGACAAAGCCAGAGCCCCCTAAGCTCCAGCTCTCTCCCAAGAGCCCCCAGAGCCCA	2403
2345	DB	GAGACAAAGCCAGAGCCCCCTAAGCTCCAGCTCTCTCCCAAGAGCCCCCAGAGCCCA	2404
2404	QY	AACAAGTCTCTCCCTTGACTTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCCCTTACCT	2463
2405	DB	AACAAGTCTCTCCCTTGACTTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCCCTTACCT	2464
2464	QY	CTGTCTCTCTGGAGCAGGCCCTTGCCAGAAACCCACAGCTCACAGAGCGGAAGCAACC	2523
2465	DB	CTGTCTCTCTGGAGCAGGCCCTTGCCAGAAACCCACAGCTCACAGAGCGGAAGCAACC	2524
2524	QY	GTCCCGGAGCAGGAACCTGCAGCAGCTGGAAATAGAATTATTCCTCAACGCTGTGCCAG	2583
2525	DB	GTCCCGGAGCAGGAACCTGCAGCAGCTGGAAATAGAATTATTCCTCAACGCTGTGCCAG	2584
2584	QY	CCATTTTCTGTGGAGGACGAGGACCAATTCCTCTCGCTCAGCATCGACGCTCTCC	2643
2585	DB	CCATTTTCTGTGGAGGACGAGGACCAATTCCTCTCGCTCAGCATCGACGCTCTCC	2644
2644	QY	CTGTCCGATGACGTGTAGAAACCCCATCAAGGCCCTCTCAAGCTCGCGGACACCCCTG	2703
2645	DB	CTGTCCGATGACGTGTAGAAACCCCATCAAGGCCCTCTCAAGCTCGCGGACACCCCTG	2704
2704	QY	AGCTCAGCGTACACTCTGTGGAGCAGCCAGGCCGAGGCTCGAAGCTCCAGCTGGAAATG	2763
2705	DB	AGCTCAGCGTACACTCTGTGGAGCAGCCAGGCCGAGGCTCGAAGCTCCAGCTGGAAATG	2764
2764	QY	GTGCTGGCCCCGGGGCGGCCACCGACACCCACGACTATTTTCATGTGTGAAGTCCAA	2823
2765	DB	GTGCTGGCCCCGGGGCGGCCACCGACACCCACGACTATTTTCATGTGTGAAGTCCAA	2824


```
QY 2824 ATACAGTCTCTTAATGGTGAACACCTGACATCCGGAGATTCCACCGGGTCAAGTGGGA 2883
DB 2825 ATACAGTCTCTTAATGGTGAACACCTGACATCCGGAGATTCCACCGGGTCAAGTGGGA 2884
QY 2884 GACATCGGCACCTGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGTGTACACAAA 2943
DB 2885 GACATCGGCACCTGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGTGTACACAAA 2944
QY 2944 GACGGGACGCTGTTTCGTACGACATGAGGTGCGAGCTCGGGCATCGACCTGCAGTGC 3003
DB 2945 GACGGGACGCTGTTTCGTACGACATGAGGTGCGAGCTCGGGCATCGACCTGCAGTGC 3004
QY 3004 ACATGCGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAGATGCGCAGCTGGAG 3063
DB 3005 ACATGCGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAGATGCGCAGCTGGAG 3064
QY 3064 AACAGGCCCTTAACCCCTGCCCTCCACCGCCGCTCCACACTCCCGGAAAGCAGCTTCCTG 3123
DB 3065 AACAGGCCCTTAACCCCTGCCCTCCACCGCCGCTCCACACTCCCGGAAAGCAGCTTCCTG 3124
QY 3124 CTCGGTGCAGATGCTGCCCTGAAACACAGGCTCAGCCGTTCCCGAGGGATTCAGGCC 3183
DB 3125 CTCGGTGCAGATGCTGCCCTGAAACACAGGCTCAGCCGTTCCCGAGGGATTCAGGCC 3184
QY 3184 CCCGGCTCAAGTGGGAAACAGGCTTCGAGCAGCAAGGTGGGGCAAGCAATGCC 3243
DB 3185 CCCGGCTCAAGTGGGAAACAGGCTTCGAGCAGCAAGGTGGGGCAAGCAATGCC 3244
QY 3244 TCCAGGATTTTACACCTGAGCCCTGCCACCTGCTGAAACACATCCGCCACGTTGA 3303
DB 3245 TCCAGGATTTTACACCTGAGCCCTGCCACCTGCTGAAACACATCCGCCACGTTGA 3304
QY 3304 GAGACAGAGGAGTGCAGGAGTTACTGGGGAAACAAACAGGGATCTTTTCTGCC 3363
DB 3305 GAGACAGAGGAGTGCAGGAGTTACTGGGGAAACAAACAGGGATCTTTTCTGCC 3364
QY 3364 CCTGCTCAGTGCAGTTCAGCTGACCCCTGGATCAGTGACCAATTTGTTGGCAGACAGG 3423
DB 3365 CCTGCTCAGTGCAGTTCAGCTGACCCCTGGATCAGTGACCAATTTGTTGGCAGACAGG 3424
QY 3424 GGAGAGCAGCTTCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTCGCCCTCACCCCTC 3483
DB 3425 GGAGAGCAGCTTCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTCGCCCTCACCCCTC 3484
QY 3484 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGCCCAAACTCAGTTCAGTGCAGAACCA 3543
DB 3485 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGCCCAAACTCAGTTCAGTGCAGAACCA 3544
QY 3544 GGTGAGCAGGTATGCCCGCCGTAGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCCTC 3603
DB 3545 GGTGAGCAGGTATGCCCGCCGTAGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCCTC 3604
QY 3604 ACCTGGCAGCTCACCCCTTTTGGGTGAGGGGAAAGAAATGCTGTGACCTCGGAAAGCT 3663
DB 3605 ACCTGGCAGCTCACCCCTTTTGGGTGAGGGGAAAGAAATGCTGTGACCTCGGAAAGCT 3664
QY 3664 CCCTGGTAGAATACACCACTTTTCAAGTTGTTGCAACACAGCTCCTGAGTTGACCTCT 3723
DB 3665 CCCTGGTAGAATACACCACTTTTCAAGTTGTTGCAACACAGCTCCTGAGTTGACCTCT 3724
QY 3724 GGTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGGTTCTCAGTCCCGACACATGT 3783
DB 3725 GGTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGGTTCTCAGTCCCGACACATGT 3784
QY 3784 GCGCCCTTGTGCTGGCTACCACTCTTCCCGCAGAGCAGAGCCCGAGCCCTTCAGGC 3843
DB 3785 GCGCCCTTGTGCTGGCTACCACTCTTCCCGCAGAGCAGAGCCCGAGCCCTTCAGGC 3844
QY 3844 CCAGCACTGCCCCAGACTCGCTGSCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTGT 3903
DB 3845 CCAGCACTGCCCCAGACTCGCTGSCACTCAGTTCCCTCATCTGTGTAAAGGTGAAGGGTGT 3904
```

```
QY 3904 GCAGGATATGCTCAGCAGGAACACTCTGTGATGGACATGATCAGTCTAGGAAAGCAG 3963
DB 3905 GCAGGATATGCTCAGCAGGAACACTCTGTGATGGACATGATCAGTCTAGGAAAGCAG 3964
QY 3964 CAGAGAGAGAGCTCCGGCGCCCGCAGCCCACTATCAGTGTCCAGCTGCTGTTCCCCAG 4023
DB 3965 CAGAGAGAGAGCTCCGGCGCCCGCAGCCCACTATCAGTGTCCAGGCTGCTGTTCCCCAG 4024
QY 4024 AGCAAGCTCAGCATCACTGACACTCACCTCCCTGCCCTTGGCCCTGGCCAGAGGTAATGC 4083
DB 4025 AGCAAGCTCAGCATCACTGACACTCACCTCCCTGCCCTTGGCCCTGGCCAGAGGTAATGC 4084
QY 4084 CGACGGCACTTTGACACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTGGCAGGGC 4143
DB 4085 CGACGGCACTTTGACACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTGGCAGGGC 4144
QY 4144 AGGCGAGGGAGTGCACACTGTAGGAGCATAGCAAGCAGGAGATGGGTGAAGGACACA 4203
DB 4145 AGGCGAGGGAGTGCACACTGTAGGAGCATAGCAAGCAGGAGATGGGTGAAGGACACA 4204
QY 4204 GTCTTGAGCTGTCCACATGATGTGACTCTCTCAAACTCTTCCAGATTTCTTAAGAATA 4263
DB 4205 GTCTTGAGCTGTCCACATGATGTGACTCTCTCAAACTCTTCCAGATTTCTTAAGAATA 4264
QY 4264 GCACCCCTTCCCCTTGGCCCTGAGCTTCTTCCAGGGGAGCTACTCAGGACTC 4323
DB 4265 GCACCCCTTCCCCTTGGCCCTGAGCTTCTTCCAGGGGAGCTACTCAGGACTC 4324
QY 4324 ACCTAGCATTAATCACTGTGATCTGAGGGGGTGTCTGCTAGCTCAACCTCTCTGGG 4383
DB 4325 ACCTAGCATTAATCACTGTGATCTGAGGGGGTGTCTGCTAGCTCAACCTCTCTGGG 4384
QY 4384 GCAGGGACGCGGAGACTCCGTTGGGAGAGCTCAATTCCTCCACATCTTGCCTAAGACAGCCTT 4443
DB 4385 GCAGGGACGCGGAGACTCCGTTGGGAGAGCTCAATTCCTCCACATCTTGCCTAAGACAGCCTT 4444
QY 4444 TGTCCAGCTGTCCACATGATGATGCTGTCTCCGGGGAGAGAGCCCGGCCCGCCAGCAC 4503
DB 4445 TGTCCAGCTGTCCACATGATGATGCTGTCTCCGGGGAGAGAGCCCGGCCCGCCAGCAC 4504
QY 4504 ATAAAGAACTGCGACCTTGGTACTGACAGCTCTGGGTGTAGAGAACTCTTTGTAAAGCAA 4563
DB 4505 ATAAAGAACTGCGACCTTGGTACTGACAGCTCTGGGTGTAGAGAACTCTTTGTAAAGCAA 4564
QY 4564 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4595
DB 4565 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4596
```

RESULT 4

PCT-US02-18947-975
; Sequence 975, Application PC/TUS0218947
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-228
; CURRENT APPLICATION NUMBER: PCT/US02/18947
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 975
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_003954
; DATABASE ENTRY DATE: 2001-06-18
PCT-US02-18947-975

Query Match 99.6%; Score 4578.4; DB 1; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 5 GGGGACTGTGCGGTGGAACTGTAGCTGTGA-AGTGGACTCTGTACCATTTGAGG 63
Db 5 GGGGACTGTGCGGTGGAACTGTAGCTGTGTAGAGGTGGACTCTGTGTACCATTTGAGG 64
QY 64 ATGTTTGGAGGATGATGTTGTGGCAGAGGCACATAAACAGGACGAGACCCCTTTGCC 123
Db 65 ATGTTTGGAGGATGATGTTGTGGCAGAGGCACATAAACAGGACGAGACCCCTTTGCC 124
QY 124 CTTGCTTTCTCCCAACCCAGGCTGACCTGTGTTCTCCAGGTCTGGGATTTCTAAGT 183
Db 125 CTTGCTTTCTCCCAACCCAGGCTGACCTGTGTTCTCCAGGTCTGGGATTTCTAAGT 184
QY 184 GACCTGCTGTGTTGTTCTCTCAGGATGAGCACAAAGCTGGGAGATGCGAGTGATG 243
Db 185 GACCTGCTGTGTTGTTCTCTCAGGATGAGCACAAAGCTGGGAGATGCGAGTGATG 244
QY 244 GAAATGGCTTGCCCAAGTGCCTCGCTCAGCAGTGGGCGACAGAAAGAACTCCCAAG 303
Db 245 GAAATGGCTTGCCCAAGTGCCTCGCTCAGCAGTGGGCGACAGAAAGAACTCCCAAG 304
QY 304 CCAAGGAGAGACCGCGCACTGGGGAAGAAACAGAGCTCGTCTACAGCTTGAGGCC 363
Db 305 CCAAGGAGAGACCGCGCACTGGGGAAGAAACAGAGCTCGTCTACAGCTTGAGGCC 364
QY 364 GTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGACGTGATTAACCAAG 423
Db 365 GTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGACGTGATTAACCAAG 424
QY 424 GGCACAGCCAAAGAGGCTCCAGGACGGCCAGCTGCCATCTCTATCATCGCCAGGCT 483
Db 425 GGCACAGCCAAAGAGGCTCCAGGACGGCCAGCTGCCATCTCTATCATCGCCAGGCT 484
QY 484 GAGTGTGAAATAGCCAAAGTTCAGCCCACTTTTCAGAACGCATTTTCATGCTGGG 543
Db 485 GAGTGTGAAATAGCCAAAGTTCAGCCCACTTTTCAGAACGCATTTTCATGCTGGG 544
QY 544 TCCAAACAGTACAGCAGTCCAGAGTCTTGATCAGATCCCAACAATGTGGCCCATGCT 603
Db 545 TCCAAACAGTACAGCAGTCCAGAGTCTTGATCAGATCCCAACAATGTGGCCCATGCT 604
QY 604 ACAGAGGGCAAAATGGCCGTGTGTTGGAAAGGAAAGCGTCGACGAAAGCCCGGAAG 663
Db 605 ACAGAGGGCAAAATGGCCGTGTGTTGGAAAGGAAAGCGTCGACGAAAGCCCGGAAG 664
QY 664 AAACGGAAGAAAGAGCTCAAAGTCCCTGGCTCATGACGAGTGGCCCTTGCCCAACCC 723
Db 665 AAACGGAAGAAAGAGCTCAAAGTCCCTGGCTCATGACGAGTGGCCCTTGCCCAACCC 724
QY 724 CTCCCCAGGACCCCTGAGCAGGAGCTGCACCATCCAGTCAGGAGGATGAGTCTCCA 783
Db 725 CTCCCCAGGACCCCTGAGCAGGAGCTGCACCATCCAGTCAGGAGGATGAGTCTCCA 784
QY 784 CTCCGCGCCCAATATGTTAGAAACACCCCGCAGTTTCAACAAGCTCTGAAGGAACAGGC 843
Db 785 CTCCGCGCCCAATATGTTAGAAACACCCCGCAGTTTCAACAAGCTCTGAAGGAACAGGC 844
QY 844 CTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGGCCGGCTCTGCTCGATCA 903
Db 845 CTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGGCCGGCTCTGCTCGATCA 904
QY 904 GAACTCCCAAACTGATCAGCCCTTGGCAATGCTGAAACAGTGTGAAACTGACACCAC 963
Db 905 GAACTCCCAAACTGATCAGCCCTTGGCAATGCTGAAACAGTGTGAAACTGACACCAC 964
QY 964 CCCAGGAGCGAGGCCCTGCGCCCTGCCCACGCAACCCCTTCCCTATAGCAGATGCT 1023
Db 965 CCCAGGAGCGAGGCCCTGCGCCCTGCCCACGCAACCCCTTCCCTATAGCAGATGCT 1024
QY 1024 CATCCCTTCCCAATTCACCCCTCTCAGGCCCTGGAACCTCACCCTCTGGAGTCTTCTG 1083
Db 1025 CATCCCTTCCCAATTCACCCCTCTCAGGCCCTGGAACCTCACCCTCTGGAGTCTTCTG 1084

QY 1084 GGCAAAATGGGCTGTGTAGACGACGAAAAACCTTTGCTGTGACCCACACCTGAGCAAACTG 1143
Db 1085 GGCAAAATGGGCTGTGTAGACGACGAAAAACCTTTGCTGTGACCCACACCTGAGCAAACTG 1144
QY 1144 GCCTGTGTAGACAGTCCAAAAGCCCTGCTGGCCCAACACCTGGAGCCAGCTGCTGTCT 1203
Db 1145 GCCTGTGTAGACAGTCCAAAAGCCCTGCTGGCCCAACACCTGGAGCCAGCTGCTGTCT 1204
QY 1204 CGTGGTGGCCATGAGAAAGTCTTCTGTGGAGAAATACCTAGTGTGCTCTCAAGGACG 1263
Db 1205 CGTGGTGGCCATGAGAAAGTCTTCTGTGGAGAAATACCTAGTGTGCTCTCAAGGACG 1264
QY 1264 GTGAGCTCAAGCCAGGCCCCACAGCTGACCAAGCTTGGCAAGACCTGGGCGAGCAGGGCC 1323
Db 1265 GTGAGCTCAAGCCAGGCCCCACAGCTGACCAAGCTTGGCAAGACCTGGGCGAGCAGGGCC 1324
QY 1324 TCCAGATCCCGGAGCCAGCCCCAAACTGAGGACAAACGAGGGTGTCTGCTCACTGAG 1383
Db 1325 TCCAGATCCCGGAGCCAGCCCCAAACTGAGGACAAACGAGGGTGTCTGCTCACTGAG 1384
QY 1384 AAACTCAAGCCAGTGGATTATGATPACCGAGAAAGTCCACTGGGCCACCGCACCAGCTC 1443
Db 1385 AAACTCAAGCCAGTGGATTATGATPACCGAGAAAGTCCACTGGGCCACCGCACCAGCTC 1444
QY 1444 CGCCTGGGCGAGAGGCTCCTTCGGAGAGGTGACAGGATGGAGGACAAAGCAGACTGGCTTC 1503
Db 1445 CGCCTGGGCGAGAGGCTCCTTCGGAGAGGTGACAGGATGGAGGACAAAGCAGACTGGCTTC 1504
QY 1504 CAGTGCCTGTCAAAGAGTGGCTGGAAGTATTTTCGGGCGAGGAGCTGATGCGATGT 1563
Db 1505 CAGTGCCTGTCAAAGAGTGGCTGGAAGTATTTTCGGGCGAGGAGCTGATGCGATGT 1564
QY 1564 GCAGGATGACTCTCACCCAGAAATTTGCTCCCTTTGATGGAGCTGTGAGAAAGGGCTTGG 1623
Db 1565 GCAGGATGACTCTCACCCAGAAATTTGCTCCCTTTGATGGAGCTGTGAGAAAGGGCTTGG 1624
QY 1624 GTCAACATCTTTCATGGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTGCTCAAGAGAGCAG 1683
Db 1625 GTCAACATCTTTCATGGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTGCTCAAGAGAGCAG 1684
QY 1684 GGTGTGTCTCCAGAGGACCGGGCCCTGTACTAAGTGGGCCAGGGCTGTGGAGGCTTGAA 1743
Db 1685 GGTGTGTCTCCAGAGGACCGGGCCCTGTACTAAGTGGGCCAGGGCTGTGGAGGCTTGAA 1744
QY 1744 TACCTCCACTCAAGAGGATTCGATGGGAGCGTCAAAGCTGCAACGCTGCTCTGCTCC 1803
Db 1745 TACCTCCACTCAAGAGGATTCGATGGGAGCGTCAAAGCTGCAACGCTGCTCTGCTCC 1804
QY 1804 AGCGATGGGAGCCAGCGAGCCCTCTGTGACTTTTGGCCATGCTGTGTCTTCAACCTGAT 1863
Db 1805 AGCGATGGGAGCCAGCGAGCCCTCTGTGACTTTTGGCCATGCTGTGTCTTCAACCTGAT 1864
QY 1864 GGCCTGGGAAAGTCTTGTCTCAGAGGGAATTACATCCCTGGCAGAGAGCCACATGGCT 1923
Db 1865 GGCCTGGGAAAGTCTTGTCTCAGAGGGAATTACATCCCTGGCAGAGAGCCACATGGCT 1924
QY 1924 CCGGAGTGGTGTGGGCGAGGAGTGCAGCGCAAGTGGATGTCTGGAGCAGCTGCTGT 1983
Db 1925 CCGGAGTGGTGTGGGCGAGGAGTGCAGCGCAAGTGGATGTCTGGAGCAGCTGCTGT 1984
QY 1984 ATGATGCTGCATGCTCAAAGCTGCGCCCTGAGCTCAAGCTTCTCCGAGGGCCGCTC 2043
Db 1985 ATGATGCTGCATGCTCAAAGCTGCGCCCTGAGCTCAAGCTTCTCCGAGGGCCGCTC 2044
QY 2044 TGCCTCAAGATTCAGGAGGCTCCGCTGTGAGGAGATCCACCCCTCTGCGGCCCT 2103
Db 2045 TGCCTCAAGATTCAGGAGGCTCCGCTGTGAGGAGATCCACCCCTCTGCGGCCCT 2104
QY 2104 CTCACAGCCCGAGGCTCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGCA 2163
Db 2105 CTCACAGCCCGAGGCTCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGCA 2164
QY 2164 GCGGAGCTGGGAGGAGGTGAACCGGGCACTTACAGCAAGTGGGAGGTGTGAAGAGCCCT 2223

Db 2165 |||||CGGAGCTGGAGGGGAGGTGAACGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT||| 2224
Qy 2224 TGGAGGGGAGGAATATAAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATTAACCAACAG 2283
Db 2225 TGGAGGGGAGGAATATAAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATTAACCAACAG 2284
Qy 2284 ACCCTCCATGCCAGCCGAGAGAGCTTTCGCAAGAGGGCCCGAGGGCCCGGCGAGCTGAG 2343
Db 2285 ACCCTCCATGCCAGCCGAGAGAGCTTTCGCAAGAGGGCCCGAGGGCCCGGCGAGCTGAG 2344
Qy 2344 GAGACACAGCAGCAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAGAGCCA 2403
Db 2345 GAGACACAGCAGCAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAGAGCCA 2404
Qy 2404 AACAAAGTCTCCTCCCTTGACTTTGAGCAAGAGAGAGTCTGGGATGTGGGAACCCCTTACCT 2463
Db 2405 AACAAAGTCTCCTCCCTTGACTTTGAGCAAGAGAGAGTCTGGGATGTGGGAACCCCTTACCT 2464
Qy 2464 CTGTCTCTCCCTGGAGCCAGCCCTGCGCAGAAACCCAGCTCAACAGAGCGGAAAGCAAC 2523
Db 2465 CTGTCTCTCCCTGGAGCCAGCCCTGCGCAGAAACCCAGCTCAACAGAGCGGAAAGCAAC 2524
Qy 2524 GTCCGGAGCAGGAACCTGACAGCTGGAATAGAAATTAATTCCTCAACAGAGCTGTCCAG 2583
Db 2525 GTCCGGAGCAGGAACCTGACAGCTGGAATAGAAATTAATTCCTCAACAGAGCTGTCCAG 2584
Qy 2584 CCAATTTCTCTGGAGGAGCAGGAGCAAAATCTCTGTCGCTCAGCATCGACAGCTCTCC 2643
Db 2585 CCAATTTCTCTGGAGGAGCAGGAGCAAAATCTCTGTCGCTCAGCATCGACAGCTCTCC 2644
Qy 2644 CTGTCCGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGGCTCGCGGGAACACCTG 2703
Db 2645 CTGTCCGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGGCTCGCGGGAACACCTG 2704
Qy 2704 AGCTCAGGCTGACTCTCTGAGCAGCAGGAGCCAGGCTCGAAGCTCAGCTGGAACATG 2763
Db 2705 AGCTCAGGCTGACTCTCTGAGCAGCAGGAGCCAGGCTCGAAGCTCAGCTGGAACATG 2764
Qy 2764 GTGTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTTGTAAGTCCAA 2823
Db 2765 GTGTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTTGTAAGTCCAA 2824
Qy 2824 ATACAGTCTCTTAATGTTGAACACCTGACATCCGGAGTTCCACCGGGTCAAAAGTGGGA 2883
Db 2825 ATACAGTCTCTTAATGTTGAACACCTGACATCCGGAGTTCCACCGGGTCAAAAGTGGGA 2884
Qy 2884 GACATCGCCTGAGCAGCAGATCCAGATCCAGCTGAGCTTTCAGCTTGGTCAACAAA 2943
Db 2885 GACATCGCCTGAGCAGCAGATCCAGATCCAGCTGAGCTTTCAGCTTGGTCAACAAA 2944
Qy 2944 GACGGGAGCTGTTGCTACGATGAGGTGCGAGCTCGGAGTCCAGCTGAGCTGAGTGC 3003
Db 2945 GACGGGAGCTGTTGCTACGATGAGGTGCGAGTCCAGCTGAGCTGAGCTGAGTGC 3004
Qy 3004 ACATGCGCCCTGATGAGCTTCCGCTGAGCTGAGGTGAGGCTCAAGCATGCGCAGCTGGAG 3063
Db 3005 ACATGCGCCCTGATGAGCTTCCGCTGAGCTGAGGTGAGGCTCAAGCATGCGCAGCTGGAG 3064
Qy 3064 AACAGGCCCTAAACCTGCTCCACCGCGCTCCACATGCTCGGGAAGCAGCTTCTCTG 3123
Db 3065 AACAGGCCCTAAACCTGCTCCACCGCGCTCCACATGCTCGGGAAGCAGCTTCTCTG 3124
Qy 3124 CTGGTGCAGATGCTGCTGAAACACAGGCTCAGGCTTCCAGGGGATTCGACAGCC 3183
Db 3125 CTGGTGCAGATGCTGCTGAAACACAGGCTCAGGCTTCCAGGGGATTCGACAGCC 3184
Qy 3184 CCCGGCTCAGTGGGAAACAGGCTCGCAGCAGCAAGTGGGGGCAAGCAATGCC 3243
Db 3185 CCCGGCTCAGTGGGAAACAGGCTCGCAGCAGCAAGTGGGGGCAAGCAATGCC 3244
Qy 3244 TCCAGGATTTACACCTGAGCCCTGCCCCACCTGTCTGTAAGAAACATCCGCCACGTGAA 3303

Db 3245 TCCAGGATTTACACCTGAGCCCTGCCCCACCTGTCTGTAAGAAACATCCGCCACGTGAA 3304
Qy 3304 GAGACAGAGGAGGATGGCAGGAGTTACTCTGGGAAACAAACAGGGATCTTTTCTGCC 3363
Db 3305 GAGACAGAGGAGGATGGCAGGAGTTACTCTGGGAAACAAACAGGGATCTTTTCTGCC 3364
Qy 3364 CCTGCTCAGTTCAGTTGGCTGACCCGCTTGGATCAGTGAACATTTGTTGGCAGACAGG 3423
Db 3365 CCTGCTCAGTTCAGTTGGCTGACCCGCTTGGATCAGTGAACATTTGTTGGCAGACAGG 3424
Qy 3424 GGAGAGCAGCTTTCAGGCTTGGGTCAGAGGGGTGGGAGCCCTTGGGCCCTCACCCCTC 3483
Db 3425 GGAGAGCAGCTTTCAGGCTTGGGTCAGAGGGGTGGGAGCCCTTGGGCCCTCACCCCTC 3484
Qy 3484 CAGGCTCTGTGAGAGTGTCAAGTGTGAAGGGGCCCAAACTCAGTTCAGTTCAGTGCAGAACCA 3543
Db 3485 CAGGCTCTGTGAGAGTGTCAAGTGTGAAGGGGCCCAAACTCAGTTCAGTTCAGTGCAGAACCA 3544
Qy 3544 GGTTCAGCAGGTATGCCCGCCGCTAGGTTAAGGGGGCCCTCTAAACCCCTTGGCTGGCCCTC 3603
Db 3545 GGTTCAGCAGGTATGCCCGCCGCTAGGTTAAGGGGGCCCTCTAAACCCCTTGGCTGGCCCTC 3604
Qy 3604 ACCTGGCAGCTCACCCCTTTTGGGTGTAGGGGAAAGATGCTGACCCCTGGGAAAGCT 3663
Db 3605 ACCTGGCAGCTCACCCCTTTTGGGTGTAGGGGAAAGATGCTGACCCCTGGGAAAGCT 3664
Qy 3664 CCCTGGTAGAATACACACACTTTTTCAGGTTGTGCAACACAGGCTCTGAGTTCACCTCT 3723
Db 3665 CCCTGGTAGAATACACACACTTTTTCAGGTTGTGCAACACAGGCTCTGAGTTCACCTCT 3724
Qy 3724 GGTTCAGCAGGACCAAGAACCAAGAGGTGTGAAGTGAAGTGTTCAGTCCCAGACATGT 3783
Db 3725 GGTTCAGCAGGACCAAGAACCAAGAGGTGTGAAGTGAAGTGTTCAGTCCCAGACATGT 3784
Qy 3784 GCGCTTTGCTGCTGGCTACCACTCTTCCCAGAGCAGCAGGCCCCCGAGCCCTTCAGGC 3843
Db 3785 GCGCTTTGCTGCTGGCTACCACTCTTCCCAGAGCAGCAGGCCCCCGAGCCCTTCAGGC 3844
Qy 3844 CCAGCAGTCCCGCAGACTCGCTGGCAGTTCCTCATCTGTAAGGTGAAGGTGAT 3903
Db 3845 CCAGCAGTCCCGCAGACTCGCTGGCAGTTCCTCATCTGTAAGGTGAAGGTGAT 3904
Qy 3904 GCAGGATATCCTGACAGGAACAGTCTGTGGATGGACATGATCAGTCTAAGGAAGCAG 3963
Db 3905 GCAGGATATCCTGACAGGAACAGTCTGTGGATGGACATGATCAGTCTAAGGAAGCAG 3964
Qy 3964 CAGAGAGAGCTCGCGGCCCGAGCCCACTATCAGTTCAGGCTGCTGGTTCGCCAG 4023
Db 3965 CAGAGAGAGCTCGCGGCCCGAGCCCACTATCAGTTCAGGCTGCTGGTTCGCCAG 4024
Qy 4024 AGCAGCTCAGCATCACACTGACACTCACCTGCTGCCCTGCGCCCGCAGAGGTACTGC 4083
Db 4025 AGCAGCTCAGCATCACACTGACACTCACCTGCTGCCCTGCGCCCGCAGAGGTACTGC 4084
Qy 4084 CGACGGCACTTTGCACTCTGATGACCTCAAGACACTTTTCATGGCTGCTGCGAGGC 4143
Db 4085 CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCTGCGAGGC 4144
Qy 4144 AGGCGAGGAGTGAACATGTAGAGCATGAAGCAGCAGAGATGGGGTGAAGGACACA 4203
Db 4145 AGGCGAGGAGTGAACATGTAGAGCATGAAGCAGCAGAGATGGGGTGAAGGACACA 4204
Qy 4204 GTCTTGAGCTGCCACATGATGACTCTCTCAAACTCTCCAGATTTCTTAAGATA 4263
Db 4205 GTCTTGAGCTGCCACATGATGACTCTCTCAAACTCTCCAGATTTCTTAAGATA 4264
Qy 4264 GCACCCCTTTCCCATTTGCCAGCTTAGCTCTTCTCCAGGGGAGTACTCAGGACTC 4323
Db 4265 GCACCCCTTTCCCATTTGCCAGCTTAGCTCTTCTCCAGGGGAGTACTCAGGACTC 4324
Qy 4324 ACGTAGCATTAATCAGCTGTGAATCGTCAGGGGGTGTCTGCTAGCCCTCAACCTCTGGG 4383
Db 4325 ACGTAGCATTAATCAGCTGTGAATCGTCAGGGGGTGTCTGCTAGCCCTCAACCTCTGGG 4384

QY 4384 GCAGGGACGCGAGACTCCGTGGGAGAAAGCTCATTTCCACAATCTTGGCAAGACAGCCTT 4443
DB 4385 GCAGGGGACGCGAGACTCCGTGGGAGAAAGCTCATTTCCACAATCTTGGCAAGACAGCCTT 4444
QY 4444 TGTCAGCTGTCCACATTTAGTTCAGACTGCTCCCGGGGAGAGAGCCCGGGCCCGCAGCAC 4503
DB 4445 TGTCAGCTGTCCACATTTAGTTCAGACTGCTCCCGGGGAGAGAGCCCGGGCCCGCAGCAC 4504
QY 4504 ATAAAGAACTGCAGCCTTGGTACTCTCAGAGTCTGGGTTGTAGAGAACTCTTTGTAAAGCAA 4563
DB 4505 ATAAAGAACTGCAGCCTTGGTACTCTCAGAGTCTGGGTTGTAGAGAACTCTTTGTAAAGCAA 4564
QY 4564 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4595
DB 4565 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4596

RESULT 5
PCT-US04-01100-975
; Sequence 975, Application PC/TUS0401100
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics LLC
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-228
; CURRENT APPLICATION NUMBER: PCT/US04/01100
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: 10/342,887
; PRIOR FILING DATE: 2003-01-15
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 975
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US04-01100-975

Query Match 99.6%; Score 4578.4; DB 2; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 5 GGGGACTGTGCGGTGGAGCTGTAGCTGTGTGA-AGGTGACTCTCTTACCATTGAGG 63
DB 5 GGGGACTGTGCGGTGGAGCTGTAGCTGTGTGAAGGTGACTCTCTTACCATTGAGG 64
QY 64 ATGTTTGGAGATGATGATGTGGCAGAGGCACACATAAACAGACAGACCCCTTTGCC 123
DB 65 ATGTTTGGAGATGATGATGTGGCAGAGGCACACATAAACAGACAGACCCCTTTGCC 124
QY 124 CTTGCTTTCTCCCCAACCCCAAGGCTGACCTGTGTCTTCCAGGCTCTGGGATTTCTAAGT 183
DB 125 CTTGCTTTCTCCCCAACCCCAAGGCTGACCTGTGTCTTCCAGGCTCTGGGATTTCTAAGT 184
QY 184 GACCTGCTCTGTGTTTGTCTCTCTCAGATGAGCACAAAGCTGGGAGATGGCAGTGATG 243
DB 185 GACCTGCTCTGTGTTTGTCTCTCTCAGATGAGCACAAAGCTGGGAGATGGCAGTGATG 244
QY 244 GAAATGGCTGCCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAAGAACTCCCAAG 303
DB 245 GAAATGGCTGCCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAAGAACTCCCAAG 304
QY 304 CAAAAGGAGAAAGACGCGCCACTGGGGAGAAACAGAGCTCCGTCCTACAGCTTGAGGCC 363
DB 305 CAAAAGGAGAAAGACGCGCCACTGGGGAGAAACAGAGCTCCGTCCTACAGCTTGAGGCC 364
QY 364 GTGGAGAAAGACCCCTGTGTCTTGGGAAAGTGGGAGATCCTGAATGAAGCTGATTAACCAAG 423
DB 365 GTGGAGAAAGACCCCTGTGTCTTGGGAAAGTGGGAGATCCTGAATGAAGCTGATTAACCAAG 424
QY 424 GGCACAGCCAGGAAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATCATCGCCAGGCT 483
DB 425 GGCACAGCCAGGAAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATCATCGCCAGGCT 484
QY 484 GAGTGTAGAAATAGCCAAGATTTCAGCCCCACCTTTTCAGAACGCAATTTTCATCGCTGGG 543

DB 485 GAGTGTAGAAATAGCCAAGATTTCAGCCCCACCTTTTCAGAAAGCAATTTTCATCGCTGGG 544
QY 544 TCCAAAAGTGTACAGCAGCTCCGAGAGTCTTGTATCAGATCCCCAACAAATGTGGCCCCATGCT 603
DB 545 TCCAAAAGTGTACAGCAGCTCCGAGAGTCTTGTATCAGATCCCCAACAAATGTGGCCCCATGCT 604
QY 604 ACAGAGGCAAAATGGCCCGTGTGTGGAAAGGGAAGCGTCCGAGCAAAAGCCCGGAAG 663
DB 605 ACAGAGGCAAAATGGCCCGTGTGTGGAAAGGGAAGCGTCCGAGCAAAAGCCCGGAAG 664
QY 664 AAAACGGAAGAAAGAGCTCAAAGTCCCTGCTCATGACAGAGTGGCCTTGGCCAAAACCC 723
DB 665 AAAACGGAAGAAAGAGCTCAAAGTCCCTGCTCATGACAGAGTGGCCTTGGCCAAAACCC 724
QY 724 CTCCCAGGACCCCTGTAGCAGAGAGCTGACCATCCAGTGCAGGAGATGAGTCTCCA 783
DB 725 CTCCCAGGACCCCTGTAGCAGAGAGCTGACCATCCAGTGCAGGAGATGAGTCTCCA 784
QY 784 CTCCGCGCCCATATGTTAGAAACACACCCGCGAGTTTCAACAAGCCTCTCAAGGAACACAGC 843
DB 785 CTCCGCGCCCATATGTTAGAAACACACCCGCGAGTTTCAACAAGCCTCTCAAGGAACACAGC 844
QY 844 CTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGCGCTTACGCGCGGCTCTGCTCGATCA 903
DB 845 CTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGCGCTTACGCGCGGCTCTGCTCGATCA 904
QY 904 GAACTCCACAAACTGATCAGGCGCTTGAATGTCTGAAACCAAGCTGTGGAACCTGCACCAC 963
DB 905 GAACTCCACAAACTGATCAGGCGCTTGAATGTCTGAAACCAAGCTGTGGAACCTGCACCAC 964
QY 964 CCCAGGACGAGGCGCCCTGCGCCGACGACCCCTTCCCTATATAGCAGACTGCT 1023
DB 965 CCCAGGACGAGGCGCCCTGCGCCGACGACCCCTTCCCTATATAGCAGACTGCT 1024
QY 1024 CATCCCTTCCATTTCCACCTCTCCAGCCCTGGAAACCTCAACCTCTGGAGTCTTCTGCTG 1083
DB 1025 CATCCCTTCCATTTCCACCTCTCCAGCCCTGGAAACCTCAACCTCTGGAGTCTTCTGCTG 1084
QY 1084 GGCATACTGGCTGTGTAGACAGCAGCAAAACCTTCCCTGACCCCACTGAGCAACTG 1143
DB 1085 GGCATACTGGCTGTGTAGACAGCAGCAAAACCTTCCCTGACCCCACTGAGCAACTG 1144
QY 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGCT 1203
DB 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGCT 1204
QY 1204 CGTGTGCTCCATGAGAAAGTCTTGTGGAGAAATACCTAGTGCATGCTTGCAGGCGAGC 1263
DB 1205 CGTGTGCTCCATGAGAAAGTCTTGTGGAGAAATACCTAGTGCATGCTTGCAGGCGAGC 1264
QY 1264 GTGAGCTCAAGCAGGCCCCACAGCTTACAGCTGGCCAGAGCTTGGGAGCAAGGCGC 1323
DB 1265 GTGAGCTCAAGCAGGCCCCACAGCTTACAGCTGGCCAGAGCTTGGGAGCAAGGCGC 1324
QY 1324 TCCAGATCCCGGAGCCCCAGCCCAAACTGAGGCAACAGAGGTGCTCTGCTCACTGAG 1383
DB 1325 TCCAGATCCCGGAGCCCCAGCCCAAACTGAGGCAACAGAGGTGCTCTGCTCACTGAG 1384
QY 1384 AAATCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGGCCACCGACCAAGCTC 1443
DB 1385 AAATCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGGCCACCGACCAAGCTC 1444
QY 1444 CGCTGGGACAGAGCTCTTGGAGAGGTGCACAGGATGGAGGAACAAGAGACTGGCTTC 1503
DB 1445 CGCTGGGACAGAGCTCTTGGAGAGGTGCACAGGATGGAGGAACAAGAGACTGGCTTC 1504
QY 1504 CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTCCGGGACAGAGAGCTGATGGCATGT 1563
DB 1505 CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTCCGGGACAGAGAGCTGATGGCATGT 1564
QY 1564 GCAGGATGACCTCAACCAGAAATGTCCCTTGTATGAGCTGTGAGAGAAAGGCGCTTGG 1623

Db 1565 GCAGGATTGACCTCACCAGAAATTGTCCTTTGTATGAGAGCTGTGAGAGAGGGCCCTTGG 1624
Qy 1624 GTCAACATCTTTCATGAGAGCTGCTGAAGGTGGCTCCTTGGGCCAGCTGGTCAAGAGAGCAG 1683
Db 1625 GTCAACATCTTTCATGAGAGCTGCTGAAGAGTGGCTCCTTGGGCCAGCTGGTCAAGAGAGCAG 1684
Qy 1684 GGCTGTCTCCAGAGAGACGGGCCCTGTGTACTACTGTGGGCCAGGCCCTGGAGGGTCTGGAA 1743
Db 1685 GGCTGTCTCCAGAGAGACGGGCCCTGTGTACTACTGTGGGCCAGGCCCTGGAGGGTCTGGAA 1744
Qy 1744 TACCTCCACTCACAAGAGGATCTGATGGGAGAGCTCAAGCTGACAAAGCTGACAAAGTGTCTCTGTGCC 1803
Db 1745 TACCTCCACTCACAAGAGGATCTGATGGGAGAGCTCAAGCTGACAAAGTGTCTCTGTGCC 1804
Qy 1804 AGCGATGGAGGCCACCGCAGCCCTCTGTGACTTTTGGCCATGTCTGTGTCTTTCAACCTGAT 1863
Db 1805 AGCGATGGAGGCCACCGCAGCCCTCTGTGACTTTTGGCCATGTCTGTGTCTTTCAACCTGAT 1864
Qy 1864 GGCCTGGGAAAGTCTTGTCTCAGAGGAGCTACATCCCTTGGGCACAGAGCCACATGGCT 1923
Db 1865 GGCCTGGGAAAGTCTTGTCTCAGAGGAGCTACATCCCTTGGGCACAGAGCCACATGGCT 1924
Qy 1924 CCGGAGGTGGTGTGGGAGAGAGCTGCGACGCCAAGGTGGATGTCTGAGCAGCTGTGT 1983
Db 1925 CCGGAGGTGGTGTGGGAGAGAGCTGCGACGCCAAGGTGGATGTCTGAGCAGCTGTGT 1984
Qy 1984 ATGATGTGTCATGTCTCAACGGGTGCCACCCCTGGACTCAGTTCCTCCGAGGGCCGCTC 2043
Db 1985 ATGATGTGTCATGTCTCAACGGGTGCCACCCCTGGACTCAGTTCCTCCGAGGGCCGCTC 2044
Qy 2044 TGCCTCAAGATTGCCAGGAGCCTCCGCTGTGAGGGAGATCCACCCCTCTCTGGCCGCT 2103
Db 2045 TGCCTCAAGATTGCCAGGAGCCTCCGCTGTGAGGGAGATCCACCCCTCTCTGGCCGCT 2104
Qy 2104 CTCACAGCCAGGCCATCCAGAGGGCTGAGGAAGAGCCATCCACCGGTGTCTGCA 2163
Db 2105 CTCACAGCCAGGCCATCCAGAGGGCTGAGGAAGAGCCATCCACCGGTGTCTGCA 2164
Qy 2164 CCGGAGCTGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTCAAGAGCCCT 2223
Db 2165 CCGGAGCTGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTCAAGAGCCCT 2224
Qy 2224 TGGAGGGAGGAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATTAACCAACAG 2283
Db 2225 TGGAGGGAGGAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATTAACCAACAG 2284
Qy 2284 ACCCTCCATGCCAGCCGAGAGAGCTTTCGCCAAGGGGCCAGGGCCCGGCCAGCTGAG 2343
Db 2285 ACCCTCCATGCCAGCCGAGAGAGCTTTCGCCAAGGGGCCAGGGCCCGGCCAGCTGAG 2344
Qy 2344 GAGACAAGCAGAGCCCTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAGAGCCA 2403
Db 2345 GAGACAAGCAGAGCCCTAAGCTCCAGCTCTCTCTCCACAGAGCCCGCCAGAGCCA 2404
Qy 2404 AACAAAGTCTCTCTCTGACTTTGAGCAAGGAGAGTCTGGGATGTGGGAACCTCTACCT 2463
Db 2405 AACAAAGTCTCTCTCTGACTTTGAGCAAGGAGAGTCTGGGATGTGGGAACCTCTACCT 2464
Qy 2464 CTGTCTCTCTGGAGCCAGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2523
Db 2465 CTGTCTCTCTGGAGCCAGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2524
Qy 2524 GTCCCGGAGCAGGAAGTGCAGCAGCTGGAATAGAAATATTCTCAACAGCCTGTCCCG 2583
Db 2525 GTCCCGGAGCAGGAAGTGCAGCAGCTGGAATAGAAATATTCTCAACAGCCTGTCCCG 2584
Qy 2584 CCATTTTCTCTGGAGGAGCAGGACAAATTTCTCTGTGCTCTCAGCATCGACAGCTCTCC 2643
Db 2585 CCATTTTCTCTGGAGGAGCAGGACAAATTTCTCTGTGCTCTCAGCATCGACAGCTCTCC 2644
Qy 2644 CTGTCCGATGACGTGAGAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGACACCCCTG 2703
Db 2645 CTGTCCGATGACGTGAGAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGACACCCCTG 2704

Qy 2704 AGCTCAGGCGTACACTCCTTGAGCAGCAGGCCCGAGGCTCGAAGCTCCAGCTCGAACAATG 2763
Db 2705 AGCTCAGGCGTACACTCCTTGAGCAGCAGGCCCGAGGCTCGAAGCTCCAGCTCGAACAATG 2764
Qy 2764 GTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGTTGTAAGTCCAA 2823
Db 2765 GTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGTTGTAAGTCCAA 2824
Qy 2824 ATACAGTCTCTTAATGTGTAACACTGACATCCGGAGTTCCACCGGGTCAAGTGGGA 2883
Db 2825 ATACAGTCTCTTAATGTGTAACACTGACATCCGGAGTTCCACCGGGTCAAGTGGGA 2884
Qy 2884 GACATCCGCACTGGCATCAGCAGCAGATCCAGCTTGCGATTTGCTTGGTTCACCAAA 2943
Db 2885 GACATCCGCACTGGCATCAGCAGCAGATCCAGCTTGCGATTTGCTTGGTTCACCAAA 2944
Qy 2944 GACGGGAGCCTGTTGCTAGCAGATGAGGTGCGAGCTCGGGATCGACCTGCGAGTGC 3003
Db 2945 GACGGGAGCCTGTTGCTAGCAGATGAGGTGCGAGCTCGGGATCGACCTGCGAGTGC 3004
Qy 3004 ACATGGCCCTGATGAGCTTCCGCTGGAGCTGAGGGTCAAGCATGGCCAGCTGAG 3063
Db 3005 ACATGGCCCTGATGAGCTTCCGCTGGAGCTGAGGGTCAAGCATGGCCAGCTGAG 3064
Qy 3064 AACAGGCCCTAACCCCTGCCCTCCACCGCGGCTCCACACTGCCGAAAGCAGCTTCCCTG 3123
Db 3065 AACAGGCCCTAACCCCTGCCCTCCACCGCGGCTCCACACTGCCGAAAGCAGCTTCCCTG 3124
Qy 3124 CTGGTGCAGATGCTGCCCTGAAAAACAGGCTCAGCGTTCCAGGGGATTTGCCAGCC 3183
Db 3125 CTGGTGCAGATGCTGCCCTGAAAAACAGGCTCAGCGTTCCAGGGGATTTGCCAGCC 3184
Qy 3184 CCCGGCTCAGTGGGAAACAGGGCTCCGAGCAGCAAGTGGGGGCAAGCAAGTCC 3243
Db 3185 CCCGGCTCAGTGGGAAACAGGGCTCCGAGCAGCAAGTGGGGGCAAGCAAGTCC 3244
Qy 3244 TCCAGGATTTCAACCTGAGCCCTGCCACCCCTGTGAAAAAACAATCCGCCACGTGAA 3303
Db 3245 TCCAGGATTTCAACCTGAGCCCTGCCACCCCTGTGAAAAAACAATCCGCCACGTGAA 3304
Qy 3304 GAGACAGAGAGGATGGCAGGAGTTACTTGGGAAAAACAAACAGGGATTTTTCGTCC 3363
Db 3305 GAGACAGAGAGGATGGCAGGAGTTACTTGGGAAAAACAAACAGGGATTTTTCGTCC 3364
Qy 3364 CCTGCTCCAGTGGGCTGACCCGCTTGGATCAGTGACCATTTTGTGGCAGACAGG 3423
Db 3365 CCTGCTCCAGTGGGCTGACCCGCTTGGATCAGTGACCATTTTGTGGCAGACAGG 3424
Qy 3424 GGAGAGCAGCTTCCAGGCTGGGTGAGAGGGTGGGCGAGCCCTTCCGCCCTCACCCCTC 3483
Db 3425 GGAGAGCAGCTTCCAGGCTGGGTGAGAGGGTGGGCGAGCCCTTCCGCCCTCACCCCTC 3484
Qy 3484 CAGGCTGCTGTGAGAGTGTCAAGTGTAAAGGGCCCAAACTCAGTTCAGTGCAGAACCA 3543
Db 3485 CAGGCTGCTGTGAGAGTGTCAAGTGTAAAGGGCCCAAACTCAGTTCAGTGCAGAACCA 3544
Qy 3544 GGTGAGCAGTATGCCCGCCGTAGGTTAAGGGGCCCTCTAAACCCCTTGCCTGGCCCTC 3603
Db 3545 GGTGAGCAGTATGCCCGCCGTAGGTTAAGGGGCCCTCTAAACCCCTTGCCTGGCCCTC 3604
Qy 3604 ACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAAAGATGCTGACCTTGGGAAGGCT 3663
Db 3605 ACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAAAGATGCTGACCTTGGGAAGGCT 3664
Qy 3664 CCCTGTGAGAAATACACCACTTTTTCAGGTTTGTGCAACACAGCTCCTGAGTGTGACCTCT 3723
Db 3665 CCCTGTGAGAAATACACCACTTTTTCAGGTTTGTGCAACACAGCTCCTGAGTGTGACCTCT 3724
Qy 3724 GGTTCAGCAAGGACCAAGAGAGTGTGTAAGTGAAGTGGTTCTCAGTCCCAGACATGT 3783
Db 3725 GGTTCAGCAAGGACCAAGAGAGTGTGTAAGTGAAGTGGTTCTCAGTCCCAGACATGT 3784

3784	Qy	GCCTCTTGTGCTGGTACCACTCTTCCGAGAGCAGAGGCCCGAGCCCTTCAGGC	3843
3785	Db	GCCTCTTGTGCTGGTACCACTCTTCCGAGAGCAGAGGCCCGAGCCCTTCAGGC	3844
3844	Qy	CCAGCACTGCCCCAGACTCCGCTGGCACTCAGTTCCCTCATCTGTTAAAGGTGAAGGGTGAT	3903
3845	Db	CCAGCACTGCCCCAGACTCCGCTGGCACTCAGTTCCCTCATCTGTTAAAGGTGAAGGGTGAT	3904
3904	Qy	GCAGGATATGCTTGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTTAAGAAAGCAG	3963
3905	Db	GCAGGATATGCTTGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTTAAGAAAGCAG	3964
3964	Qy	CAGAGAGACGTCGGGGCCCCAGCCCCACTATCAGTGTCCAGCGTGTGGTTCCTCCCGAG	4023
3965	Db	CAGAGAGACGTCGGGGCCCCAGCCCCACTATCAGTGTCCAGCGTGTGGTTCCTCCCGAG	4024
4024	Qy	AGCACGCTCAGGATCACACTGCACACTCACCTTGCCCTGCCCCCTGGCCAGAGGGTACTGC	4083
4025	Db	AGCACGCTCAGGATCACACTGCACACTCACCTTGCCCTGCCCCCTGGCCAGAGGGTACTGC	4084
4084	Qy	CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTTCATATGGGTGGCCCTCTGGCAGGGC	4143
4085	Db	CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTTCATATGGGTGGCCCTCTGGCAGGGC	4144
4144	Qy	AGGCGAGGAGTGACACTGTAGGAGCATAGCAAGCCAGAGATGGGGTGAAGGGACACA	4203
4145	Db	AGGCGAGGAGTGACACTGTAGGAGCATAGCAAGCCAGAGATGGGGTGAAGGGACACA	4204
4204	Qy	GTCTTGAGCTGTCCACATGCACTGTGACTCTCTCAAACTCTTCCAGATTTTCTCTAAGAATA	4263
4205	Db	GTCTTGAGCTGTCCACATGCACTGTGACTCTCTCAAACTCTTCCAGATTTTCTCTAAGAATA	4264
4264	Qy	GCACCCCTTCCCATTTGCCCGACGTTAGCCTCTTCTCCAGGGGAGCTACTCAGAGCTC	4323
4265	Db	GCACCCCTTCCCATTTGCCCGACGTTAGCCTCTTCTCCAGGGGAGCTACTCAGAGCTC	4324
4324	Qy	ACGTAGCATTAATAACGCTGTGAATCGTCAGGGGGTGTCTGTAGCCTCAACCTCTCTGGG	4383
4325	Db	ACGTAGCATTAATAACGCTGTGAATCGTCAGGGGGTGTCTGTAGCCTCAACCTCTCTGGG	4384
4384	Qy	GCAGGGAGCCCGAGACTCCGTGGGAGAAAGCTCATTCCTCCATCTTGGCCCAAGACAGCCTT	4443
4385	Db	GCAGGGAGCCCGAGACTCCGTGGGAGAAAGCTCATTCCTCCATCTTGGCCCAAGACAGCCTT	4444
4444	Qy	TGTCACGCTGTCCACATTTGAGTCAGACTGCTCCGGGGAGAGACCCCGCCCGCCAGCAC	4503
4445	Db	TGTCACGCTGTCCACATTTGAGTCAGACTGCTCCGGGGAGAGACCCCGCCCGCCAGCAC	4504
4504	Qy	ATAAAGAACTGCAGCCTTGGTACTTGCAGAGTCTGGGTTGTAGAGAACTCTTTTGTAAAGCAA	4563
4505	Db	ATAAAGAACTGCAGCCTTGGTACTTGCAGAGTCTGGGTTGTAGAGAACTCTTTTGTAAAGCAA	4564
4564	Qy	TAAAGTTTGGGGTGATGACAAATGTTAAAAA	4595
4565	Db	TAAAGTTTGGGGTGATGACAAATGTTAAAAA	4596

RESULT 6

[illegible]

QY 1024 CATCCCTTTCCATTTCCACCCCTCTCAGGCCCTGGAAACCTCTCAACCTCTGGAATCTTCTCTG 1083
DB CATCCCTTTCCATTTCCACCCCTCTCAGGCCCTGGAAACCTCTCAACCTCTGGAATCTTCTCTG 1084
QY 1084 GGCAGAGTGGGCTGTGTAGACAGCAGAGAAACCTTGTCTGACCCACACACCTCTGAGCAACTG 1143
DB 1085 GGCAGAGTGGGCTGTGTAGACAGCAGAGAAACCTTGTCTGACCCACACACCTCTGAGCAACTG 1144
QY 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACACCTGGAGCCAGCTGCTGTCT 1203
DB 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACACCTGGAGCCAGCTGCTGTCT 1204
QY 1204 CGTGTGCCCATGAGAAATTTTCTGTGTGAGGAAATACCTAGTGTGATGCTCTCAAGGACG 1263
DB 1205 CGTGTGCCCATGAGAAATTTTCTGTGTGAGGAAATACCTAGTGTGATGCTCTCAAGGACG 1264
QY 1264 GTGAGCTCAAGCCAGGCCCAAGCCTGACCTGACAGCCTGCGCAAGACCTGGGAGCAGGGGC 1323
DB 1265 GTGAGCTCAAGCCAGGCCCAAGCCTGACAGCCTGCGCAAGACCTGGGAGCAGGGGC 1324
QY 1324 TCCAGATCCCGGAGCCAGGCCCAAACTGAGGACAAAGAGGCTGCTCTCACTGAG 1383
DB 1325 TCCAGATCCCGGAGCCAGGCCCAAACTGAGGACAAAGAGGCTGCTCTCACTGAG 1384
QY 1384 AAATCAAGCCAGTGGATTTATGATGTAACGAGAAAGTCCACTGGGCCACGACCAAGCTC 1443
DB 1385 AAATCAAGCCAGTGGATTTATGATGTAACGAGAAAGTCCACTGGGCCACGACCAAGCTC 1444
QY 1444 CGCTGGGACAGAGCTCTCTGGAGAGGTGACAGGATGAGGACAAAGACAGTGGGCTTC 1503
DB 1445 CGCTGGGACAGAGCTCTCTGGAGAGGTGACAGGATGAGGACAAAGACAGTGGGCTTC 1504
QY 1504 CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTTGGGAGAGGAGCTGATGGCATGT 1563
DB 1505 CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTTGGGAGAGGAGCTGATGGCATGT 1564
QY 1564 GCAGGATTTGACCTCAACCCAGAAATTTGCTCTTTGATGAGAGCTGTGAGAGAGGCTTTGG 1623
DB 1565 GCAGGATTTGACCTCAACCCAGAAATTTGCTCTTTGATGAGAGCTGTGAGAGAGGCTTTGG 1624
QY 1624 GTCAACATCTTCTGAGAGCTCTGGAAGGTGGCTCTCTGGGCCAGAGTGGTCAAGAGGAG 1683
DB 1625 GTCAACATCTTCTGAGAGCTCTGGAAGGTGGCTCTCTGGGCCAGAGTGGTCAAGAGGAG 1684
QY 1684 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGCTGGAGGGTCTGGA 1743
DB 1685 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGCTGGAGGGTCTGGA 1744
QY 1744 TACCTCCACTCAGAGAGGATTTCTGATGGGACCGTCAAAAGCTGACAGCTGCTCTGTCC 1803
DB 1745 TACCTCCACTCAGAGAGGATTTCTGATGGGACCGTCAAAAGCTGACAGCTGCTCTGTCC 1804
QY 1804 AGCGATGGGAGCCAGCAGCCCTCTGTGACTTTGGCCAGTGTGTGTCTTCAACCTGAT 1863
DB 1805 AGCGATGGGAGCCAGCAGCCCTCTGTGACTTTGGCCAGTGTGTGTCTTCAACCTGAT 1864
QY 1864 GGCTGTGGAAAGTCTTGTCTCAGAGGAGCTACATCCCTGGCAAGAGCCCAATGGCT 1923
DB 1865 GGCTGTGGAAAGTCTTGTCTCAGAGGAGCTACATCCCTGGCAAGAGCCCAATGGCT 1924
QY 1924 CCGAGGTGGTGTCTGGGAGGAGCTGCGACCCCAAGGTGGATGTCTGGAGCAGCTGTGT 1983
DB 1925 CCGAGGTGGTGTCTGGGAGGAGCTGCGACCCCAAGGTGGATGTCTGGAGCAGCTGTGT 1984
QY 1984 ATGATGTGACATGCTCAACGGCTGCGACCCCTGGACTCAAGTCTTCCGAGGCGCTC 2043
DB 1985 ATGATGTGACATGCTCAACGGCTGCGACCCCTGGACTCAAGTCTTCCGAGGCGCTC 2044
QY 2044 TGCCTCAAGATTGCCAGGAGCTCGGCTGTGAGGAGATCCCAACCTCTCTGGCCCT 2103
DB 2045 TGCCTCAAGATTGCCAGGAGCTCGGCTGTGAGGAGATCCCAACCTCTCTGGCCCT 2104

QY 2104 CTCAAGCCCGGCGCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGCA 2163
DB 2105 CTCAAGCCCGGCGCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGCA 2164
QY 2164 GCGAGCTGGGAGGAGGTGAAACCGGCGCATCAAGCAAGTGGGAGGTCTCAAGAGCCCT 2223
DB 2165 GCGAGCTGGGAGGAGGTGAAACCGGCGCATCAAGCAAGTGGGAGGTCTCAAGAGCCCT 2224
QY 2224 TGGAGGGGAGAAATATAAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAAC 2283
DB 2225 TGGAGGGGAGAAATATAAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAAC 2284
QY 2284 ACCCTCATGCCAGCCGAGAGAGCTTTGCGCAAGGGCCCCCAGGGCCCCCGGCGAGCTGAG 2343
DB 2285 ACCCTCATGCCAGCCGAGAGAGCTTTGCGCAAGGGCCCCCAGGGCCCCCGGCGAGCTGAG 2344
QY 2344 GAGACACAGCAGAGCCCTTAAGCTCAGGCTCTCTCTCCACAGAGGCCCCCAAGAGCCA 2403
DB 2345 GAGACACAGCAGAGCCCTTAAGCTCAGGCTCTCTCTCCACAGAGGCCCCCAAGAGCCA 2404
QY 2404 AACAAAGTCTCTCTCTCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGGGAAACCTTTACT 2463
DB 2405 AACAAAGTCTCTCTCTCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGGGAAACCTTTACT 2464
QY 2464 CTGTCTCTCTCTGGAGCAGGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAGCAAC 2523
DB 2465 CTGTCTCTCTCTGGAGCAGGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAGCAAC 2524
QY 2524 GTCCCGAGCAGGAACTGACAGCTGGAATAGAAATTTCTCTCAACAGGCTCTCCAG 2583
DB 2525 GTCCCGAGCAGGAACTGACAGCTGGAATAGAAATTTCTCTCAACAGGCTCTCCAG 2584
QY 2584 CCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTCTCAGCATCGAGGCTCTCC 2643
DB 2585 CCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTCTCAGCATCGAGGCTCTCC 2644
QY 2644 CTGTCTGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCTG 2703
DB 2645 CTGTCTGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCTG 2704
QY 2704 AGCTCAGGCTGACCTCTCTGGAGCAGCCGAGGCTCGAAGCTCGAGCTGGAACATG 2763
DB 2705 AGCTCAGGCTGACCTCTCTGGAGCAGCCGAGGCTCGAAGCTCGAGCTGGAACATG 2764
QY 2764 GTGTCTGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAAGTCCAA 2823
DB 2765 GTGTCTGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAAGTCCAA 2824
QY 2824 ATACAGTCTCTTAATGTGAACACCTGCAATCCGGAGTTTCAACCGGCTCAAGTGGGA 2883
DB 2825 ATACAGTCTCTTAATGTGAACACCTGCAATCCGGAGTTTCAACCGGCTCAAGTGGGA 2884
QY 2884 GACATCCGCACTGGCATCAGCAGCCAGATCCAGCTGAGGCTTTCAGTTGGTCAACCAA 2943
DB 2885 GACATCCGCACTGGCATCAGCAGCCAGATCCAGCTGAGGCTTTCAGTTGGTCAACCAA 2944
QY 2944 GAGCGGAGCTGTCTGCTACGATGAGGTGCGAGCTCGGAGCTCGAGCTGAGTGC 3003
DB 2945 GAGCGGAGCTGTCTGCTACGATGAGGTGCGAGCTCGGAGCTCGAGCTGAGTGC 3004
QY 3004 ACATGCGCCCTGATGTCAGCTTCCGCTGGAGCTGGAGGCTCAAGCATGGCCAGCTGGAG 3063
DB 3005 ACATGCGCCCTGATGTCAGCTTCCGCTGGAGCTGGAGGCTCAAGCATGGCCAGCTGGAG 3064
QY 3064 AACAGGCTTAACTGCTCTTCAACCGCGGCTTCACTGTCCGGAAGAGGCTTCTCTG 3123
DB 3065 AACAGGCTTAACTGCTCTTCAACCGCGGCTTCACTGTCCGGAAGAGGCTTCTCTG 3124
QY 3124 CTGGTGCAGATGCTGCTGCTGAAACACAGGCTCAGGCTTCCAGGGATTCAGGCT 3183
DB 3125 CTGGTGCAGATGCTGCTGCTGAAACACAGGCTCAGGCTTCCAGGGATTCAGGCT 3184
QY 3184 CCCCGGCTCAGTGGGAAACAGGCGCTCGCAGCAGCAAGTGGGGGCAAGCAGATGCC 3243

Db 365 GTGGAGAGAGCCCTGTGTTCTCGGAAAGTGGAGATCCTGAATGACGTGATTAACCAAG 424
Qy 424 GGCACAGCCAGGAAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATCATCGCCAGGCT 483
Db 425 GGCACAGCCAGGAAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATCATCGCCAGGCT 484
Qy 484 GAGTGTGAGATAGCCAGAGTTTCAAGCCCACTTTTTCAGAACGCAITTTTATCGCTGG 543
Db 485 GAGTGTGAGATAGCCAGAGTTTCAAGCCCACTTTTTCAGAACGCAITTTTATCGCTGG 544
Qy 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCCATGCT 603
Db 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCCATGCT 604
Qy 604 ACAGAGGCAAAATGGCCCGTGTGTTGGAGGGAAGCGTTCGAGCAAAAGCCCGGAAG 663
Db 605 ACAGAGGCAAAATGGCCCGTGTGTTGGAGGGAAGCGTTCGAGCAAAAGCCCGGAAG 664
Qy 664 AAACGGAGAGAGAGCTCAAGTCCCTGGCTCATGACAGAGTGGCTTGGCCMAACCC 723
Db 665 AAACGGAGAGAGAGCTCAAGTCCCTGGCTCATGACAGAGTGGCTTGGCCMAACCC 724
Qy 724 CTCCCGAGCCCTAGCAGAGAGTGCACCATCCAGTGCAGAGAGTGAAGTCTCCA 783
Db 725 CTCCCGAGCCCTAGCAGAGAGTGCACCATCCAGTGCAGAGAGTGAAGTCTCCA 784
Qy 784 CTCCCGAGCCCTATGTTAGAAACACCCCGCAGTTCACCAAGCCCTCTGAAGGAACCAAGC 843
Db 785 CTCCCGAGCCCTATGTTAGAAACACCCCGCAGTTCACCAAGCCCTCTGAAGGAACCAAGC 844
Qy 844 CTGGGGCAATCTGTTTAAAGCAGCTTGGGAGAGGCTTACGGCCGGCTCTGCTCGATCA 903
Db 845 CTGGGGCAATCTGTTTAAAGCAGCTTGGGAGAGGCTTACGGCCGGCTCTGCTCGATCA 904
Qy 904 GAATCCCAAACTGATCAGCCCTTGCAATGTCTGAAACCAAGTGTGAAACTGCAACAC 963
Db 905 GAATCCCAAACTGATCAGCCCTTGCAATGTCTGAAACCAAGTGTGAAACTGCAACAC 964
Qy 964 CCCCAGGAGGAGGCCCTGCCCCGACGACACCCCTTCCCTATAGCAGACTGCT 1023
Db 965 CCCCAGGAGGAGGCCCTGCCCCGACGACACCCCTTCCCTATAGCAGACTGCT 1024
Qy 1024 CATCCCTTCCATTCACCCCTCTCAGCCCTGGAAACCTCAACCTCTGGAGTCCCTTCCTG 1083
Db 1025 CATCCCTTCCATTCACCCCTCTCAGCCCTGGAAACCTCAACCTCTGGAGTCCCTTCCTG 1084
Qy 1084 GGCATACTGGCTGTGTAGACAGCAGAAACCTTGCCTGACACCACTGAGCAAACTG 1143
Db 1085 GGCATACTGGCTGTGTAGACAGCAGAAACCTTGCCTGACACCACTGAGCAAACTG 1144
Qy 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTGGAGGCCAGCTGCTGCT 1203
Db 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTGGAGGCCAGCTGCTGCT 1204
Qy 1204 CGTGTGCCATGAGAAAGTTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGAGC 1263
Db 1205 CGTGTGCCATGAGAAAGTTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGAGC 1264
Qy 1264 GTGAGCTCAAGCAGGCCCAAGCTGACAGCTGGCCAAAGACTGGGAGCAGCAGGGC 1323
Db 1265 GTGAGCTCAAGCAGGCCCAAGCTGACAGCTGGCCAAAGACTGGGAGCAGCAGGGC 1324
Qy 1324 TCCAGATCCCGGAGCCAGCCCAAACTGAGGACCAAGGAGTGTCTGCTCACTGAG 1383
Db 1325 TCCAGATCCCGGAGCCAGCCCAAACTGAGGACCAAGGAGTGTCTGCTCACTGAG 1384
Qy 1384 AAATCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCCACGACCACTGCTC 1443
Db 1385 AAATCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCCACGACCACTGCTC 1444
Qy 1444 CGCTGGCAGAGGCTCTTCGGAGAGGTGCAAGGATGGAGGCAAGCAGACTGGCTTC 1503
Db 1445 CGCTGGCAGAGGCTCTTCGGAGAGGTGCAAGGATGGAGGCAAGCAGACTGGCTTC 1504

Qy 1504 CAGTGCCTGTCAAAAAGGTCGCTCGAAGTATTTTCGGGAGAGGAGCTGATGSCATGT 1563
Db 1505 CAGTGCCTGTCAAAAAGGTCGCTCGAAGTATTTTCGGGAGAGGAGCTGATGSCATGT 1564
Qy 1564 GCAGGATTGACCTCACCCAGAAATTTGTCCTTTGTATGAGAGCTGTGAGAGAGGGCTTGG 1623
Db 1565 GCAGGATTGACCTCACCCAGAAATTTGTCCTTTGTATGAGAGCTGTGAGAGAGGGCTTGG 1624
Qy 1624 GTCAACATCTTCAATGGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTGCTCAAGGAGCAG 1683
Db 1625 GTCAACATCTTCAATGGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTGCTCAAGGAGCAG 1684
Qy 1684 GGCTGTCTCCAGAGGACCCGGGCCCTGTACTACCTGGGCCAGGGCCCTGGAGGGTCTGAA 1743
Db 1685 GGCTGTCTCCAGAGGACCCGGGCCCTGTACTACCTGGGCCAGGGCCCTGGAGGGTCTGAA 1744
Qy 1744 TACCTCCACTCACAAGGATTTCTGATGGGAGCTCAAGCTGACAACGCTGCTCTGTCC 1803
Db 1745 TACCTCCACTCACAAGGATTTCTGATGGGAGCTCAAGCTGACAACGCTGCTCTGTCC 1804
Qy 1804 AGCGATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGCTTCAACCTGAT 1863
Db 1805 AGCGATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGCTTCAACCTGAT 1864
Qy 1864 GGCTGGGAAAGTCTTGTCTCAGAGGGAGCTACATCCCTGGCACAGAGACCCACATGGCT 1923
Db 1865 GGCTGGGAAAGTCTTGTCTCAGAGGGAGCTACATCCCTGGCACAGAGACCCACATGGCT 1924
Qy 1924 CCGAGGTGTGTGTGGGAGGAGCTGCGAAGCTGGAAGTGTGTGGAGCAGCTGTGT 1983
Db 1925 CCGAGGTGTGTGTGGGAGGAGCTGCGAAGCTGGAAGTGTGTGGAGCAGCTGTGT 1984
Qy 1984 ATGATGTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTC 2043
Db 1985 ATGATGTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTC 2044
Qy 2044 TGCTTCAAGATTGCCAGGAGCTCCGCTGTGAGGGAGATCCCAACCTCTCGTGGCCCT 2103
Db 2045 TGCTTCAAGATTGCCAGGAGCTCCGCTGTGAGGGAGATCCCAACCTCTCGTGGCCCT 2104
Qy 2104 CTCAAGCCCAAGGCTTCAAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGCA 2163
Db 2105 CTCAAGCCCAAGGCTTCAAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGCA 2164
Qy 2164 GCGAGCTGGAGGGAGGTGAACCGGGCCTACAGCAAGTGGAGGCTCTGAAGAGCCCT 2223
Db 2165 GCGAGCTGGAGGGAGGTGAACCGGGCCTACAGCAAGTGGAGGCTCTGAAGAGCCCT 2224
Qy 2224 TGGAGGGAGAAATATAAGAACCAAGACATCCACCGCAATCAAGCCAAATACCCAG 2283
Db 2225 TGGAGGGAGAAATATAAGAACCAAGACATCCACCGCAATCAAGCCAAATACCCAG 2284
Qy 2284 ACCCTCATGCCAGCCGAGAGCTTTTCGCAAGGGCCCCCAGGGCCCCCGCCAGCTGAG 2343
Db 2285 ACCCTCATGCCAGCCGAGAGCTTTTCGCAAGGGCCCCCAGGGCCCCCGCCAGCTGAG 2344
Qy 2344 GAGACAAAGCAGAGGCCCTTAAGCTCAGCTCCTCTCTCCCAAGAGGCCCTCCAGAGCCA 2403
Db 2345 GAGACAAAGCAGAGGCCCTTAAGCTCAGCTCCTCTCTCCCAAGAGGCCCTCCAGAGCCA 2404
Qy 2404 AACAGTCTCTCTCCCTTGAATTTTGAAGAGGAGTCTGGGATGTGGAAACCTTACT 2463
Db 2405 AACAGTCTCTCTCCCTTGAATTTTGAAGAGGAGTCTGGGATGTGGAAACCTTACT 2464
Qy 2464 CTGTCTCTCTGGAGCCAGCCCTTGGCAGAAACCCAGCTCAACAGAGCGGAAAGCAACC 2523
Db 2465 CTGTCTCTCTGGAGCCAGCCCTTGGCAGAAACCCAGCTCAACAGAGCGGAAAGCAACC 2524
Qy 2524 GTCCCGAGCAGGAACTGACAGAGCTGGAATATTAAGATTTTCTCAACAGCCTGTCCAG 2583
Db 2525 GTCCCGAGCAGGAACTGACAGAGCTGGAATATTAAGATTTTCTCAACAGCCTGTCCAG 2584

QY 2584 CCATTTTCTCTGAGGAGCAGGAGCAAAATCTCTCGTGCCTCAGCATCGACAGCCCTCTCC 2643
Db 2585 CCATTTTCTCTGAGGAGCAGGAGCAAAATCTCTCGTGCCTCAGCATCGACAGCCCTCTCC 2644
QY 2644 CTGTCGGATGACAGTGAAGAAACCCATCAAGAGCCTCTCAAGCTCGGGGACACCCCTG 2703
Db 2645 CTGTCGGATGACAGTGAAGAAACCCATCAAGAGCCTCTCAAGCTCGGGGACACCCCTG 2704
QY 2704 AGCTCAGCGGTACATCTCTGGAGCAGCCAGCGCGAGGCTCGAAGCTTCAGCTTGGAAACATG 2763
Db 2705 AGCTCAGCGGTACATCTCTGGAGCAGCCAGCGCGAGGCTCGAAGCTTCAGCTTGGAAACATG 2764
QY 2764 GTGCTGGCCCGGGGGCGGCCACCCAGACACCCCAAGCTATTTCAATGGGTGAAGTCCAA 2823
Db 2765 GTGCTGGCCCGGGGGCGGCCACCCAGACACCCCAAGCTATTTCAATGGGTGAAGTCCAA 2824
QY 2824 ATACAGTCTCTTAATGGTGAACACCTGCAATCCGGAGTTCACACCGGGTCAAGTGGGA 2883
Db 2825 ATACAGTCTCTTAATGGTGAACACCTGCAATCCGGAGTTCACACCGGGTCAAGTGGGA 2884
QY 2884 GACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTCAGCTTGGTCAACAA 2943
Db 2885 GACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTCAGCTTGGTCAACAA 2944
QY 2944 CACGGGCAGCCTGTTCGCTTACACATGAGAGTGCAGACTCGGGCATCGACTGCAGTGC 3003
Db 2945 GACGGGCAGCCTGTTCGCTTACACATGAGAGTGCAGACTCGGGCATCGACTGCAGTGC 3004
QY 3004 ACATCGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGCTGGAG 3063
Db 3005 ACATCGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGCTGGAG 3064
QY 3064 PACAGGCCCTAACCTCGCCCTCCACCGCGGCTCCACATCGCGGAAAGCAGCTTCCTG 3123
Db 3065 AACAGGCCCTAACCTCGCCCTCCACCGCGGCTCCACATCGCGGAAAGCAGCTTCCTG 3124
QY 3124 CTCGGTGCAGATGTGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATTTGCCAGCC 3183
Db 3125 CTCGGTGCAGATGTGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATTTGCCAGCC 3184
QY 3184 CCCCCTGTCAGTGGGAAACAGGCGCTCGCAGCAGCAAGTGGGGGCAAGCAGAAATGCC 3243
Db 3185 CCCCCTGTCAGTGGGAAACAGGCGCTCGCAGCAGCAAGTGGGGGCAAGCAGAAATGCC 3244
QY 3244 TCCCAGGATTTTCAACACCTGAGCCCTGCCACCCCTGCTGAAACACATCCGCGCAGTGA 3303
Db 3245 TCCCAGGATTTTCAACACCTGAGCCCTGCCACCCCTGCTGAAACACATCCGCGCAGTGA 3304
QY 3304 GAGACAGAGGAGATGGCAGGAGTTTACCTGGGGAAACAAACAGGGATCTTTTCTGCC 3363
Db 3305 GAGACAGAGGAGATGGCAGGAGTTTACCTGGGGAAACAAACAGGGATCTTTTCTGCC 3364
QY 3364 CTGCTCCAGTCCAGTGGCCTGACCCGCTTGGATCAGTGACCATTTGTTGGCAGACAGG 3423
Db 3365 CTGCTCCAGTCCAGTGGCCTGACCCGCTTGGATCAGTGACCATTTGTTGGCAGACAGG 3424
QY 3424 GGAGAGCAGCTTCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTCGCGCCCTCACCCTC 3483
Db 3425 GGAGAGCAGCTTCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTCGCGCCCTCACCCTC 3484
QY 3484 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGGCCCAAACTCAGGTTCAGTGCAGAAACA 3543
Db 3485 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGGCCCAAACTCAGGTTCAGTGCAGAAACA 3544
QY 3544 GGTACAGAGTATGCCCGCCCTAGGTTAAGGGGCCCTCTAAACCCCTTGCCTGGCCTC 3603
Db 3545 GGTACAGAGTATGCCCGCCCTAGGTTAAGGGGCCCTCTAAACCCCTTGCCTGGCCTC 3604
QY 3604 ACCTCGCGAGCTCACCCCTTTTGGGTAGGGGAAAGAAATCCCTGACCCCTGGGAGGCT 3663
Db 3605 ACCTCGCGAGCTCACCCCTTTTGGGTAGGGGAAAGAAATCCCTGACCCCTGGGAGGCT 3664
QY 3664 CCTCGTAGAATACACCAACACTTTTTCAGGTTGTGCAACACAGGCTCCTGAGTTGACCTCT 3723

Db 3665 CCCTGGTAGAATAACACACACTTTTTCAGTTGTGCAACACAGGTCTCTGAGTTGACCTCT 3724
QY 3724 GGTTCAGCAAGGACCAAGAGGTGTAGTGAAGTGTCTCAGTCCCCAGACATGT 3783
Db 3725 GGTTCAGCAAGGACCAAGAGGTGTAGTGAAGTGTCTCAGTCCCCAGACATGT 3784
QY 3784 GCCCTTTGCTGTGGCTTACCCTCTTCCAGAGCAGCAGGCCCCAGGCCCTTTCAGGC 3843
Db 3785 GCCCTTTGCTGTGGCTTACCCTCTTCCAGAGCAGCAGGCCCCAGGCCCTTTCAGGC 3844
QY 3844 CCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCTCAATCTGTAAAGGTGAAGGTTGAT 3903
Db 3845 CCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCTCAATCTGTAAAGGTGAAGGTTGAT 3904
QY 3904 GCAGGATATGCTGCACAGGAACAGTCTGTGATGGACATGATCAGTGTCTAGGAAAGCAG 3963
Db 3905 GCAGGATATGCTGCACAGGAACAGTCTGTGATGGACATGATCAGTGTCTAGGAAAGCAG 3964
QY 3964 CAGAGAGAGAGCTCCGGGCCCCAGGCCCACTATCAGTGTCCAGCGTCTGTTCCCGAG 4023
Db 3965 CAGAGAGAGAGCTCCGGGCCCCAGGCCCACTATCAGTGTCCAGCGTCTGTTCCCGAG 4024
QY 4024 AGCACAGCTCAGCATCACACTGACACTCACCTGCCCCCTGGCCAGAGGGTACTGTC 4083
Db 4025 AGCACAGCTCAGCATCACACTGACACTCACCTGCCCCCTGGCCAGAGGGTACTGTC 4084
QY 4084 CGACGGCACTTTGCACTCTGATGACCTCAAGACATTTTCATGGCTGCCCTCTGGCAGGGC 4143
Db 4085 CGACGGCACTTTGCACTCTGATGACCTCAAGACATTTTCATGGCTGCCCTCTGGCAGGGC 4144
QY 4144 AGGGCAGGGCAGTGCACACTGTAGGAGCATAGCAAGCCAGAGGATGGGGTGAAGGACACA 4203
Db 4145 AGGGCAGGGCAGTGCACACTGTAGGAGCATAGCAAGCCAGAGGATGGGGTGAAGGACACA 4204
QY 4204 GTCTTGAGCTGTCCAATGCAATGTGACTCTTCAAACCTTCTTCCAGATTTCTTAAGAATA 4263
Db 4205 GTCTTGAGCTGTCCAATGCAATGTGACTCTTCAAACCTTCTTCCAGATTTCTTAAGAATA 4264
QY 4264 GCACCCCTTCCCATGGCCCCAGGTTAGCCTTCTCCAGGGGAGCTACTCAGGACTC 4323
Db 4265 GCACCCCTTCCCATGGCCCCAGGTTAGCCTTCTTCTCCAGGGGAGCTACTCAGGACTC 4324
QY 4324 ACGTAGCATTAATCAGCTGTGAATCGTCAGGGGTGTCTGCTAGCCTCAACCTCTGGG 4383
Db 4325 ACGTAGCATTAATCAGCTGTGAATCGTCAGGGGTGTCTGCTAGCCTCAACCTCTGGG 4384
QY 4384 GCAGGGGACGCGCAGACTCCCGTGGGAGAAAGCTCATTTCCACATCTTGGCAGACAGCCTT 4443
Db 4385 GCAGGGGACGCGCAGACTCCCGTGGGAGAAAGCTCATTTCCACATCTTGGCAGACAGCCTT 4444
QY 4444 TGTCCAGCTGTCCCATGAGTTCAGACTGTCTCCGGGGAGAGAGCCCCGGCCCCCAGCAC 4503
Db 4445 TGTCCAGCTGTCCCATGAGTTCAGACTGTCTCCGGGGAGAGAGCCCCGGCCCCCAGCAC 4504
QY 4504 ATAAAGAACTGCAGCCTTGGTACTGACAGTCTGGGTGTAGAGAACTCTTTTGAAGCAA 4563
Db 4505 ATAAAGAACTGCAGCCTTGGTACTGACAGTCTGGGTGTAGAGAACTCTTTTGAAGCAA 4564
QY 4564 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4595
Db 4565 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4596

RESULT 8
US-09-981-397A-17
; Sequence 17, Application US/09981397A
; GENERAL INFORMATION:
; APPLICANT: Axxima Pharmaceuticals AG
; APPLICANT: Schubart, Daniel
; APPLICANT: Habenberger, Peter
; APPLICANT: Stein-Gerlach, Matthias
; APPLICANT: Bevec, Dorian

; TITLE OF INVENTION: Cellular Kinases Involved in Cytomegalovirus Infection and their
; TITLE OF INVENTION: Inhibition
; FILE REFERENCE: AXM-004.1 US
; CURRENT APPLICATION NUMBER: US/09/981,397A
; CURRENT FILING DATE: 2002-06-28
; PRIOR FILING DATE: 60/240,750
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-981-397A-17

Query Match 99.6%; Score 4578.4; DB 45; Length 4596;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY	5	GGGGACTGTGCGGTGTGGAACGTGTAGCTGTGGA-AGGTGACCTCTGTTACCAATTGAGG	63
DB	5	GGGGACTGTGCGGTGTGGAACGTGTAGCTGTGGAAGGTGACTCTGTTACCAATTGAGG	64
QY	64	ATGTTTGGAGGATGAGTATGTGTGCGAGAGGACACATAAACAGCAGACACCTTTGCC	123
DB	65	ATGTTTGGAGGATGAGTATGTGTGCGAGAGGACACATAAACAGCAGACACCTTTGCC	124
QY	124	CCTGCTTTCTCCCAACCAAGGCTGACCTGTGTTCTCCAGGTCTGGGANTCTAAGT	183
DB	125	CCTGCTTTCTCCCAACCAAGGCTGACCTGTGTTCTCCAGGTCTGGGANTCTAAGT	184
QY	184	GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAGCTGGGAGATGGCAGTGATG	243
DB	185	GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAGCTGGGAGATGGCAGTGATG	244
QY	244	GAATGCGCTGCGCAGGTGCGCTGCTGAGAGTGGGGGAGAGAGAACTCCCAAG	303
DB	245	GAATGCGCTGCGCAGGTGCGCTGCTGAGAGTGGGGGAGAGAGAACTCCCAAG	304
QY	304	CCAAAGGAGAGAGCGCCCACTGGGGAAGAAACAGAGCTCCGCTACAAGCTTGAGGCC	363
DB	305	CCAAAGGAGAGAGCGCCCACTGGGGAAGAAACAGAGCTCCGCTACAAGCTTGAGGCC	364
QY	364	GTGAGAGAGAGCCCTGTGTTCTGCGGAAAGTGGAGATCCTGAATGACGTGATTACCAAG	423
DB	365	GTGAGAGAGAGCCCTGTGTTCTGCGGAAAGTGGAGATCCTGAATGACGTGATTACCAAG	424
QY	424	GGCAGAGCAAGAGAGGTCCGAGGAGGCGAGCTGCCATCTCTATCATCGCCAGGCT	483
DB	425	GGCAGAGCAAGAGAGGTCCGAGGAGGCGAGCTGCCATCTCTATCATCGCCAGGCT	484
QY	484	GAGTGTGAGATAGCCAGAGTTGAGCCCACTTTTCAGAACGATTTTCATCGCTGGG	543
DB	485	GAGTGTGAGATAGCCAGAGTTGAGCCCACTTTTCAGAACGATTTTCATCGCTGGG	544
QY	544	TCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATGCT	603
DB	545	TCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATGCT	604
QY	604	ACAGAGGCAAAATGGCCCGTGTGTTGGAGGGAAGCGTCCGAGCAAAAGCCCGGAAG	663
DB	605	ACAGAGGCAAAATGGCCCGTGTGTTGGAGGGAAGCGTCCGAGCAAAAGCCCGGAAG	664
QY	664	AAAAGGAGAGAGAGCTCAAAGTCCCTGGCTCATGAGAGTGGCCCTTGCCCAACCC	723
DB	665	AAAAGGAGAGAGAGCTCAAAGTCCCTGGCTCATGAGAGTGGCCCTTGCCCAACCC	724
QY	724	CTCCCAAGAGAGAGCTGAGCAGAGCTGACCACTCCAGTGCAGGAGTGTCTCCA	783
DB	725	CTCCCAAGAGAGAGCTGAGCAGAGCTGACCACTCCAGTGCAGGAGTGTCTCCA	784
QY	784	CTCCGCGGCCCATATGTTAGAAACACCCCGAGTTTCAACAGCTCTTGAAGAAACCAAGGC	843

DB	785	CTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTACCAAGCCTCTCTGAAGAACCAAGGC	844
QY	844	CTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTTACGGCGGCTCTCTCCTCGATCA	903
DB	845	CTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTTACGGCGGCTCTCTCCTCGATCA	904
QY	904	GAACCTCCACAAACTGATCAGGCCCCCTTGAATGTCTGAACACAGCTGTGGAACCTGACACAC	963
DB	905	GAACCTCCACAAACTGATCAGGCCCCCTTGAATGTCTGAACACCAAGCTGTGGAACCTGACACAC	964
QY	964	CCCCAGGAGGAGGCCCCCTTGCCTCCACGACACCCCTTCCCTATATAGCAGACTGCT	1023
DB	965	CCCCAGGAGGAGGCCCCCTTGCCTCCACGACACCCCTTCCCTATATAGCAGACTGCT	1024
QY	1024	CATCCCTTCCCATTTCCACCTCTCTCAGCCTCTGGAACCTCTCCTCTTGGAGTCTCTTCTG	1083
DB	1025	CATCCCTTCCCATTTCCACCTCTCTCAGCCTCTGGAACCTCTCCTCTTGGAGTCTCTTCTG	1084
QY	1084	GGCAAACTGGCCTGTGTAGACAGCCAGAAACCTTGTCTGTGACCCACACCTGAGCAAACTG	1143
DB	1085	GGCAAACTGGCCTGTGTAGACAGCCAGAAACCTTGTCTGTGACCCACACCTGAGCAAACTG	1144
QY	1144	GCCTGTGTAGACAGTCCAAAGCCCCCTTGCCTGGCCCCACCTTGGAGCCAGCTGCTGTCT	1203
DB	1145	GCCTGTGTAGACAGTCCAAAGCCCCCTTGCCTGGCCCCACCTTGGAGCCAGCTGCTGTCT	1204
QY	1204	CGTGGTCCCATGAGAAAGTTTCTGTGGAGAAATACTAGTGCATGCTCTGCAAGCGAGC	1263
DB	1205	CGTGGTCCCATGAGAAAGTTTCTGTGGAGAAATACTAGTGCATGCTCTGCAAGCGAGC	1264
QY	1264	GTGAGCTCAAGCCAGGCCCCACAGCTGACAGCTTGGCCTGCAAGAGCGGCGC	1323
DB	1265	GTGAGCTCAAGCCAGGCCCCACAGCTGACAGCTTGGCCTGCAAGAGCGGCGC	1324
QY	1324	TCCAGATCCCGGGAGCCCCCAAACTGAGGACACAGGGGTGCTCTGCTCACTGAG	1383
DB	1325	TCCAGATCCCGGGAGCCCCCAAACTGAGGACACAGGGGTGCTCTGCTCACTGAG	1384
QY	1384	AAACTCAAGCCAGTGGATTTATGAGTACCGAGAAAGTCCACTGGGCGCACGACCAAGCTC	1443
DB	1385	AAACTCAAGCCAGTGGATTTATGAGTACCGAGAAAGTCCACTGGGCGCACGACCAAGCTC	1444
QY	1444	CGCTGGGCGCAGAGCTCTCTCGGAGAGTGCAAGGATGGAGGACAAGCAGACTGGCTTC	1503
DB	1445	CGCTGGGCGCAGAGCTCTCTCGGAGAGTGCAAGGATGGAGGACAAGCAGACTGGCTTC	1504
QY	1504	CAGTGGCTGTCAAAAGGTGGCTGGAAGTATTTGGGCGAGAGGAGCTGATGSCATGT	1563
DB	1505	CAGTGGCTGTCAAAAGGTGGCTGGAAGTATTTGGGCGAGAGGAGCTGATGSCATGT	1564
QY	1564	GCAGATTGACCTCACCCAGAAATGTCCTTTGTATGAGAGCTGTGAGAGAGGGCTTTGG	1623
DB	1565	GCAGATTGACCTCACCCAGAAATGTCCTTTGTATGAGAGCTGTGAGAGAGGGCTTTGG	1624
QY	1624	GTCAACATCTTTCTGAGAGTGGCTTCCCTGGGCGCAGCTGGTCAAGGAGCAG	1683
DB	1625	GTCAACATCTTTCTGAGAGTGGCTTCCCTGGGCGCAGCTGGTCAAGGAGCAG	1684
QY	1684	GGCTGTCTCCAGAGGACCGGCGCTGTATCTAAGCTGGGCGCAGGCTTGGAGGGTCTGAA	1743
DB	1685	GGCTGTCTCCAGAGGACCGGCGCTGTATCTAAGCTGGGCGCAGGCTTGGAGGGTCTGAA	1744
QY	1744	TACCTCCACTCACAAAGGATTTGATGGGAGCTCAAGCTGACAGAGCTCTCTGCTCC	1803
DB	1745	TACCTCCACTCACAAAGGATTTGATGGGAGCTCAAGCTGACAGAGCTCTCTGCTCC	1804
QY	1804	AGCATGGGAGCAGCAGGCGCTCTGTGATTTTGGCCATGCTGTGTCTTCAACCTGAT	1863
DB	1805	AGCATGGGAGCAGCAGGCGCTCTGTGATTTTGGCCATGCTGTGTCTTCAACCTGAT	1864
QY	1864	GGCTGGGAAAGTCTTGTCTCAGAGGAGTATCACTCCCTGGCAGAGAGCCACATGGCT	1923
DB	1865	GGCTGGGAAAGTCTTGTCTCAGAGGAGTATCACTCCCTGGCAGAGAGCCACATGGCT	1924

Qy	1924	CCGAGAGTGTGTCTGGGCAGGAGCTGCGACGCGCAAGGTGGATGTCTGGAGCAGCTGCTGT	1983
Db	1925	CCGAGAGTGTGTCTGGGCAGGAGCTGCGACGCGCAAGGTGGATGTCTGGAGCAGCTGCTGT	1984
Qy	1984	ATGATGCTGCACATGCTCAACGGGTGCCACCCCTGGACTCAGTTCTTCCAGAGGCGCGCTC	2043
Db	1985	ATGATGCTGCACATGCTCAACGGGTGCCACCCCTGGACTCAGTTCTTCCAGAGGCGCGCTC	2044
Qy	2044	TGCCTCAAGATTGCGACGAGCCTCCGCTGTGAGGGAGATCCACACCTCCTCGGCCCT	2103
Db	2045	TGCCTCAAGATTGCGACGAGCCTCCGCTGTGAGGGAGATCCACACCTCCTCGGCCCT	2104
Qy	2104	CTCACGCCCAGGCGCATCAAGAGGGCTGAGAAAGAGGCCATCCACCGCTGTCTGCA	2163
Db	2105	CTCACGCCCAGGCGCATCAAGAGGGCTGAGAAAGAGGCCATCCACCGCTGTCTGCA	2164
Qy	2164	GCGGAGCTGGAGGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT	2223
Db	2165	GCGGAGCTGGAGGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT	2224
Qy	2224	TGGAGGGAGNATATAAGAACCAAGACATCCACCGCCAAATCAAGCCATTTACCACG	2283
Db	2225	TGGAGGGAGNATATAAGAACCAAGACATCCACCGCCAAATCAAGCCATTTACCACG	2284
Qy	2284	ACCTCTCATGCGCCAGCAGAGAGCTTTTCGCAAGGGCCCCAGGGCCCCCGGCAGCTGAG	2343
Db	2285	ACCTCTCATGCGCCAGCAGAGAGCTTTTCGCAAGGGCCCCCGGCAGCTGAG	2344
Qy	2344	GAGACACAGCAGAGCCCCCTAAGCTCCAGCCTCTCTCCACACAGAGCCCCCAGAGCCA	2403
Db	2345	GAGACACAGCAGAGCCCCCTAAGCTCCAGCCTCTCTCCACACAGAGCCCCCAGAGCCA	2404
Qy	2404	AACAGTCTCTCTCTTGACTTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCTTTACCT	2463
Db	2405	AACAGTCTCTCTCTTGACTTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCTTTACCT	2464
Qy	2464	CTGTCTCTCTGGAGCAGGCCCTTCGCAGAAAAACCCAGCTCACACAGAGCGGAAGCAACC	2523
Db	2465	CTGTCTCTCTGGAGCAGGCCCTTCGCAGAAAAACCCAGCTCACACAGAGCGGAAGCAACC	2524
Qy	2524	GTCCCGGACAGGAACTGCAGCAGCTGGAAATAGAAATTTCTCTCAACAGCCTGTCCCAG	2583
Db	2525	GTCCCGGACAGGAACTGCAGCAGCTGGAAATAGAAATTTCTCTCAACAGCCTGTCCCAG	2584
Qy	2584	CCATTTTCTCTGGAGGACGAGGACAAATTCTCTGTGCTCTCAGCATCGACGCTCTCC	2643
Db	2585	CCATTTTCTCTGGAGGACGAGGACAAATTCTCTGTGCTCTCAGCATCGACGCTCTCC	2644
Qy	2644	CTGTGCGATGACGTGAGAGAACCCATCAAGGCCCTCAAGAGCTCGCGGGACCCCTG	2703
Db	2645	CTGTGCGATGACGTGAGAGAACCCATCAAGGCCCTCAAGAGCTCGCGGGACCCCTG	2704
Qy	2704	AGCTCAGCGGTACACTCTCTGGAGCAGCCAGGCCGAGGCTCGAAGCTCCAGCTCGAAACATG	2763
Db	2705	AGCTCAGCGGTACACTCTCTGGAGCAGCCAGGCCGAGGCTCGAAGCTCCAGCTCGAAACATG	2764
Qy	2764	GTGCTGGCCCCGGGGCGGCCACCGCAACCCCAAGCTATTTCATGTGTGAAGTCCAA	2823
Db	2765	GTGCTGGCCCCGGGGCGGCCACCGCAACCCCAAGCTATTTCATGTGTGAAGTCCAA	2824
Qy	2824	ATACAGTCTCTTAATGTGAACCACTGCACATCCGGAGTTCCACCGGTCGAAAGTGGGA	2883
Db	2825	ATACAGTCTCTTAATGTGAACCACTGCACATCCGGAGTTCCACCGGTCGAAAGTGGGA	2884
Qy	2884	GACATCGCACTGGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGCTTGGTCAACAAA	2943
Db	2885	GACATCGCACTGGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGCTTGGTCAACAAA	2944
Qy	2944	GACGGGAGCCTGTTCGTACGACATGGAGGTGCCAGCTCGGGCATCGACCTGCAGTGC	3003
Db	2945	GACGGGAGCCTGTTCGTACGACATGGAGGTGCCAGCTCGGGCATCGACCTGCAGTGC	3004

Qy	3004	ACA	CTGGCCCTTGATGGCAGCTTCGCGCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAG	3063					
Db	3005	ACA	CTGGCCCTTGATGGCAGCTTCGCGCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAG	3064					
Qy	3064	ACA	GCGGCCCTAAACCCCTGCGCCCTCCACCGCGGGCTCCACACTGCCGGAAGACAGACCTTCCTG	3123					
Db	3065	ACA	GCGGCCCTAAACCCCTGCGCCCTCCACCGCGGGCTCCACACTGCCGGAAGACAGACCTTCCTG	3124					
Qy	3124	CT	CGGTGCAGATGCTGCGCCCTGAAAAACA	CAGGCTCAGCGCGTTC	CC	CAGGGGATTC	CAGGCC	3183	
Db	3125	CT	CGGTGCAGATGCTGCGCCCTGAAAAACA	CAGGCTCAGCGCGTTC	CC	CAGGGGATTC	CAGGCC	3184	
Qy	3184	CCC	CGGCTCA	CAGTGGGAA	CC	AGGGCTCG	CACAGCAGCAAGTGGGGCAAGCAGATGCC	3243	
Db	3185	CCC	CGGCTCA	CAGTGGGAA	CC	AGGGCTCG	CACAGCAGCAAGTGGGGCAAGCAGATGCC	3244	
Qy	3244	TCC	CAGGATTTACA	CTTGAGCCCTGCGCCCA	CCCTGCTG	AAAAAACA	CATCGGCCACGTGA	3303	
Db	3245	TCC	CAGGATTTACA	CTTGAGCCCTGCGCCCA	CCCTGCTG	AAAAAACA	CATCGGCCACGTGA	3304	
Qy	3304	GAG	CAGAAGGAGATGGCAGGATTTAC	CTGGGGAA	CA	AAAA	CAGGATCTTTTTCTGCC	3363	
Db	3305	GAG	CAGAAGGAGATGGCAGGATTTAC	CTGGGGAA	CA	AAAA	CAGGATCTTTTTCTGCC	3364	
Qy	3364	CCT	GCTCCAGT	CTGAGTTGG	CCTG	ACCCGCTTGGAT	CAGTGACCATTTGTTGGCAGACAGG	3423	
Db	3365	CCT	GCTCCAGT	CTGAGTTGG	CCTG	ACCCGCTTGGAT	CAGTGACCATTTGTTGGCAGACAGG	3424	
Qy	3424	GGA	GACAGCTTC	CAGCTGGGTG	CAGAAGGGGTGGGGC	AGCCCTT	CGGCCCTC	CACGCTC	3483
Db	3425	GGA	GACAGCTTC	CAGCTGGGTG	CAGAAGGGGTGGGGC	AGCCCTT	CGGCCCTC	CACGCTC	3484
Qy	3484	CAG	CTGCTG	TAGAGTG	TCAAGTG	TGAAGGCC	CAAACTCAGGTT	CAGTGAGCAACCA	3543
Db	3485	CAG	CTGCTG	TAGAGTG	TCAAGTG	TGAAGGCC	CAAACTCAGGTT	CAGTGAGCAACCA	3544
Qy	3544	GGT	CAGCAGAT	TATGCCCG	CCGTAGGT	TAAAGGGGGCCCT	CTAAACCCCTTGGCTGGCCCTC	3603	
Db	3545	GGT	CAGCAGAT	TATGCCCG	CCGTAGGT	TAAAGGGGGCCCT	CTAAACCCCTTGGCTGGCCCTC	3604	
Qy	3604	AC	CTGGCAG	CTCA	CCCTTTTGGGTG	TAGGGGAAAGAA	TGCTGACCCCTGGGAAGGCT	3663	
Db	3605	AC	CTGGCAG	CTCA	CCCTTTTGGGTG	TAGGGGAAAGAA	TGCTGACCCCTGGGAAGGCT	3664	
Qy	3664	CC	CTGTAG	GAATACA	CCACATTTT	CAGGTTG	TGCAACACAGGTC	CTGAGTTGACCTCT	3723
Db	3665	CC	CTGTAG	GAATACA	CCACATTTT	CAGGTTG	TGCAACACAGGTC	CTGAGTTGACCTCT	3724
Qy	3724	GGT	TTCAGC	CAAGCA	CAAAAGAG	GTGTGAAGTGAAGTGGT	TTCTCAGTCCCCAGACATGT	3783	
Db	3725	GGT	TTCAGC	CAAGCA	CAAAAGAG	GTGTGAAGTGAAGTGGT	TTCTCAGTCCCCAGACATGT	3784	
Qy	3784	G	CCCTTTTGT	GCTGGCT	ACCACTCTT	CC	CAGACAGCAGGCCCGACCCCTTCAGGC	3843	
Db	3785	G	CCCTTTTGT	GCTGGCT	ACCACTCTT	CC	CAGACAGCAGGCCCGACCCCTTCAGGC	3844	
Qy	3844	CA	GAC	CTGCCCCAC	AGACTCGCTGGCACTCAGTTT	CCCTCATCTGTAAAGGTGAAGGTTGAT	3903		
Db	3845	CA	GAC	CTGCCCCAC	AGACTCGCTGGCACTCAGTTT	CCCTCATCTGTAAAGGTGAAGGTTGAT	3904		
Qy	3904	GC	AGGATAT	GCCTTGA	CAGGAA	CAGTCTGTGGATGGACATCATCAGTGT	CAAGGAAAGCAG	3963	
Db	3905	GC	AGGATAT	GCCTTGA	CAGGAA	CAGTCTGTGGATGGACATCATCAGTGT	CAAGGAAAGCAG	3964	
Qy	3964	CAG	AGAGAG	AGCTCG	CGCGCC	CCACAGCCCACTATCAGTGT	CCAGCGTGTGTTCCCCCAG	4023	
Db	3965	CAG	AGAGAG	AGCTCG	CGCGCC	CCACAGCCCACTATCAGTGT	CCAGCGTGTGTTCCCCCAG	4024	
Qy	4024	AGA	CAC	GCTCAG	ATACACTG	ACATCAG	CCCTGCCCCCTGGCCACAGGGTACTGC	4083	
Db	4025	AGA	CAC	GCTCAG	ATACACTG	ACATCAG	CCCTGCCCCCTGGCCACAGGGTACTGC	4084	
Qy	4084	CG	AC	GGCAC	TTTGC	ACTCTCAT	GACCTCAAGACATTTTCATGGCTG	CGCCCTCTGGCAGGGC	4143

```
Db 4085 CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTATGGCTGCGCTCTGGCAGGCG 4144
Qy 4144 AGGCGAGGCACTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGTGAAGGACACA 4203
Db 4145 AGGCGAGGCACTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGTGAAGGACACA 4204
Qy 4204 GTCTTTGAGCTGCCACATGATGATGACCTCTCAAACTCTTCCAGATTCTCTAAGAATA 4263
Db 4205 GTCTTTGAGCTGCCACATGATGATGACCTCTCAAACTCTTCCAGATTCTCTAAGAATA 4264
Qy 4264 GCACCCCTCTCCCACTGCGCCAGCTTAGCCTCTTCTCCAGGGGAGCTACTCAGGACTC 4323
Db 4265 GCACCCCTCTCCCACTGCGCCAGCTTAGCCTCTTCTCCAGGGGAGCTACTCAGGACTC 4324
Qy 4324 ACGTAGCATTAATCAGCTGTGAATCGTCAAGGGGTGTCTGCTAGCCTCAACCTCTCTGGG 4383
Db 4325 ACGTAGCATTAATCAGCTGTGAATCGTCAAGGGGTGTCTGCTAGCCTCAACCTCTCTGGG 4384
Qy 4384 CGAGGGGACGCGAGACTCCGTGGGAGAGCTCATTCGCAATCTTGCCAAAGACAGCCTT 4443
Db 4385 GCAGGGGACGCGAGACTCCGTGGGAGAGAGCTCATTCGCAATCTTGCCAAAGACAGCCTT 4444
Qy 4444 TGTCCAGCTGTCCCACTTGAATGCTGAGCTGCTCCCGGGGAGAGAGCCCGGCCCCCAGCAC 4503
Db 4445 TGTCCAGCTGTCCCACTTGAATGCTGAGCTGCTCCCGGGGAGAGAGCCCGGCCCCCAGCAC 4504
Qy 4504 ATAAAGAACTGCAGCTTGGTGTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTGTAAGCAA 4563
Db 4505 ATAAAGAACTGCAGCTTGGTGTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTGTAAGCAA 4564
Qy 4564 TAAAGTTGGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4595
Db 4565 TAAAGTTGGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4596
```

RESULT 9

```
US-10-172-118-975
; Sequence 975, Application US/10172118
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 975
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_003954
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-975
```

```
Query Match 99.6%; Score 4578.4; DB 49; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 5 GGGGAGCTGTCCGCTGTGGAACGTGTAGCTGTGA-AGGTGACTCTGTTACCACTTGGG 63
Db 5 GGGGAGCTGTCCGCTGTGGAACGTGTAGCTGTGAGAGGTGGAAGTCTGTTTACCACTTGGG 64
Qy 64 ATGTTTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 123
```

```
Db 65 ATGTTTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 124
Qy 124 CCTGCTTTCTCCCAACCAAGGCTGACCTGTGTTCTCCAGGCTCTGGGATTTCTAAGT 183
Db 125 CCTGCTTTCTCCCAACCAAGGCTGACCTGTGTTCTCCAGGCTCTGGGATTTCTAAGT 184
Qy 184 GACCTGCTCTGTGTTGTTGTTCTCTCAGGATGAGCAAAAGCTGGAGATGGCAGTGTATG 243
Db 185 GACCTGCTCTGTGTTGTTGTTCTCTCAGGATGAGCAAAAGCTGGAGATGGCAGTGTATG 244
Qy 244 GAAATGGCTTCCCAAGGCTGGCTGAGCAAGCTGGGAGAGCAAGGAACTCCCCAAG 303
Db 245 GAAATGGCTTCCCAAGGCTGGCTGAGCAAGCTGGGAGAGCAAGGAACTCCCCAAG 304
Qy 304 CCAAAGGAGAAAGACGCGCCCACTGGGGAAGAAAACAGAGCTCCGTCACAAGCTTCAGAGCC 363
Db 305 CCAAAGGAGAAAGACGCGCCCACTGGGGAAGAAAACAGAGCTCCGTCACAAGCTTCAGAGCC 364
Qy 364 GTGAGAAAGAGCCCTGTGTTCTCGGAAAGTGGGAGATTCCTGAATGACGTTGATACCAAG 423
Db 365 GTGAGAAAGAGCCCTGTGTTCTCGGAAAGTGGGAGATTCCTGAATGACGTTGATACCAAG 424
Qy 424 GGCAAGCAAGAAAGGCTCCGAGGAGGAGGAGCTGAGCTGCTATCTATCATGCCCCAGGCT 483
Db 425 GGCAAGCAAGAAAGGCTCCGAGGAGGAGGAGCTGAGCTGCTATCTATCATGCCCCAGGCT 484
Qy 484 GAGTGTGAGAAATACCAAGAGTTTCAAGCCCACTTTTTCAGAAACGATTTTTCATCGCTGG 543
Db 485 GAGTGTGAGAAATACCAAGAGTTTCAAGCCCACTTTTTCAGAAACGATTTTTCATCGCTGG 544
Qy 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCCATGCT 603
Db 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCCATGCT 604
Qy 604 ACAGAGGCAAAATGGCCCGTGTGTTGGAGGAAAGCGTCGAGCAAAAGCCCGGGAAG 663
Db 605 ACAGAGGCAAAATGGCCCGTGTGTTGGAGGAAAGCGTCGAGCAAAAGCCCGGGAAG 664
Qy 664 AACGGAAGAGAGAGCTCAAGTCCCTGCTGCTCATGCAAGAGTGGCTTGGCCCAAAACCC 723
Db 665 AACGGAAGAGAGAGCTCAAGTCCCTGCTGCTCATGCAAGAGTGGCTTGGCCCAAAACCC 724
Qy 724 CTCGCCAGAGACCCCTGAGCAGGAGAGCTGACCACTCCAGTCCAGGAGGATGAGTCTCCA 783
Db 725 CTCGCCAGAGACCCCTGAGCAGGAGAGCTGACCACTCCAGTCCAGGAGGATGAGTCTCCA 784
Qy 784 CTCGCCAGAGACCCCTGAGCAGGAGAGCAACCCCGAGTTTCAAGGCTCTGGAAGAAACAGGC 843
Db 785 CTCGCCAGAGACCCCTGAGCAGGAGAGCAACCCCGAGTTTCAAGGCTCTGGAAGAAACAGGC 844
Qy 844 CTTGGGCAACTCTGTTTAAAGCAGCTTGGGAGGAGGCTTCCGCGGCTCTGCTCGATCA 903
Db 845 CTTGGGCAACTCTGTTTAAAGCAGCTTGGGAGGAGGCTTCCGCGGCTCTGCTCGATCA 904
Qy 904 GAACTCCACAAACTGATCAGCCCTTCAATGTTCTGAACCACTGGAAGAACTGCACCAC 963
Db 905 GAACTCCACAAACTGATCAGCCCTTCAATGTTCTGAACCACTGGAAGAACTGCACCAC 964
Qy 964 CCCCAGAGGAGGCCCCCTGCCCCCTGCCCAAGCAACCCCTTCCCTATAGCAGACTGCT 1023
Db 965 CCCCAGAGGAGGCCCCCTGCCCCCTGCCCAAGCAACCCCTTCCCTATAGCAGACTGCT 1024
Qy 1024 CATCCCTTCCGATTCACCCCTCTCCAGCCCTGGGAAACCTCAACCTCTGGAGTCTCTCTG 1083
Db 1025 CATCCCTTCCGATTCACCCCTCTCCAGCCCTGGGAAACCTCAACCTCTGGAGTCTCTCTG 1084
Qy 1084 GGCACAACTGCGCTGTGTAGACAGCAGCAAAACCCCTTGGCTGACCCACACCTGAGCAAACTG 1143
Db 1085 GGCACAACTGCGCTGTGTAGACAGCAGCAAAACCCCTTGGCTGACCCACACCTGAGCAAACTG 1144
Qy 1144 GCCTGTGTAGACAGTCCAAAGCCCTTCCCTGGCCCAACCTTGGAGCCAGCTGCTGTCT 1203
```


Db 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACACCTTGGAGCCCAAGCTGCTGTCT 1204
Qy 1204 CGTGTGCCCATGAGAAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGCAGC 1263
Db 1205 CGTGTGCCCATGAGAAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGCAGC 1264
Qy 1264 GTGAGCTCAAGCCAGGCCCCACAGCCCTGACCAGCCCTGGCCAAAGACCTGGGCAAGCACGGGGC 1323
Db 1265 GTGAGCTCAAGCCAGGCCCCACAGCCCTGACCAGCCCTGGCCAAAGACCTGGGCAAGCACGGGGC 1324
Qy 1324 TCCAGATCCCGGAGCCAGCCCCAAAACCTAGGACAACAGAGGGTGTCTGCTCACTGAG 1383
Db 1325 TCCAGATCCCGGAGCCAGCCCCAAAACCTAGGACAACAGAGGGTGTCTGCTCACTGAG 1384
Qy 1384 AAACCTCAAGCCAGTGGATATACAGTACCGAGAAGATGCCACTGGGCCACGACACAGCTC 1443
Db 1385 AAACCTCAAGCCAGTGGATATACAGTACCGAGAAGATGCCACTGGGCCACGACACAGCTC 1444
Qy 1444 CGCCTGGGCAGAGGCTCCTTCCGAGAGAGTGCAACAGGATGGAGCAAAAGCAGACTGGCTTC 1503
Db 1445 CGCCTGGGCAGAGGCTCCTTCCGAGAGAGTGCAACAGGATGGAGCAAAAGCAGACTGGCTTC 1504
Qy 1504 CAGTGGCTGTCAAAAAGTGCGCTCGGCTGGAAGTATTTCCGGCAGAGAGCTGATGGCATGT 1563
Db 1505 CAGTGGCTGTCAAAAAGTGCGCTCGGCTGGAAGTATTTCCGGCAGAGAGCTGATGGCATGT 1564
Qy 1564 GCAGGATTGACCTCACCCAGAAATGTCCCTTTGTATGGAGCTGTGAGAGAAGGGCTTGG 1623
Db 1565 GCAGGATTGACCTCACCCAGAAATGTCCCTTTGTATGGAGCTGTGAGAGAAGGGCTTGG 1624
Qy 1624 GTCAACATCTTCATGAGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAG 1683
Db 1625 GTCAACATCTTCATGAGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAG 1684
Qy 1684 GGCTGTCTCCAGAGAGACCGGGCCCTGTACTACCTGGGCCAGGGCCCTGGAGGGTCTGGAA 1743
Db 1685 GGCTGTCTCCAGAGAGACCGGGCCCTGTACTACCTGGGCCAGGGCCCTGGAGGGTCTGGAA 1744
Qy 1744 TACCTCCATCAGAGAGATTCGATGGGAGCGTCAAGCTGACAAAGCTGCTCCTGTC 1803
Db 1745 TACCTCCATCAGAGAGATTCGATGGGAGCGTCAAGCTGACAAAGCTGCTCCTGTC 1804
Qy 1804 AGCGATGGAGGACGAGCCCTCTGTGACTTTGGCCATGCTGTGTGTCTTCAACCTGAT 1863
Db 1805 AGCGATGGAGGACGAGCCCTCTGTGACTTTGGCCATGCTGTGTGTCTTCAACCTGAT 1864
Qy 1864 GGCTGGGAAAATCTCTTGTCTCAAGGGACTACATCCCTGGCAAGAGACCCACATGGCT 1923
Db 1865 GGCTGGGAAAATCTCTTGTCTCAAGGGACTACATCCCTGGCAAGAGACCCACATGGCT 1924
Qy 1924 CCGGAGGTGGTCTGGGCAAGAGCTCGACGCGCAAGGTGGATGTCTGGAGCAGCTGCTGT 1983
Db 1925 CCGGAGGTGGTCTGGGCAAGAGCTCGACGCGCAAGGTGGATGTCTGGAGCAGCTGCTGT 1984
Qy 1984 ATGATGCTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTCTTCCGAGGGCCGCTC 2043
Db 1985 ATGATGCTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTCTTCCGAGGGCCGCTC 2044
Qy 2044 TGCCTCAAGATTGCCAGAGGCTCGCTGTGAGGGAGATCCCAACCTCTCTGGGCCCT 2103
Db 2045 TGCCTCAAGATTGCCAGAGGCTCGCTGTGAGGGAGATCCCAACCTCTCTGGGCCCT 2104
Qy 2104 CTCACAGCCAGGCCATCCAGAGGGCTGAGGAAGAGCCCATCCACCGCTGTGCA 2163
Db 2105 CTCACAGCCAGGCCATCCAGAGGGCTGAGGAAGAGCCCATCCACCGCTGTGCA 2164
Qy 2164 CGGAGCTGGAGGGAAGGTGAACCGGCACTACAGCAAGTGGAGGTCTGAAGAGCCCT 2223
Db 2165 CGGAGCTGGAGGGAAGGTGAACCGGCACTACAGCAAGTGGAGGTCTGAAGAGCCCT 2224
Qy 2224 TGGAGGGAGATATATAAGAAACAAGACATCAACCGCCAAATCAAGCCAAATACCAACAG 2283
Db 2225 TGGAGGGAGATATATAAGAAACAAGACATCAACCGCCAAATCAAGCCAAATACCAACAG 2284

Qy 2284 ACCCTCCATGCCAGCCGAGAGAGCTTTGCGCAAGGGCCCCAGGGCCCCGGCCAGCTGAG 2343
Db 2285 ACCCTCCATGCCAGCCGAGAGAGCTTTGCGCAAGGGCCCCAGGGCCCCGGCCAGCTGAG 2344
Qy 2344 GAGCAACAGGAGAGAGCCCTTAAGCTCAGGCTCTCTCTCCACAGAGCCGCCAGAGCA 2403
Db 2345 GAGCAACAGGAGAGAGCCCTTAAGCTCAGGCTCTCTCTCCACAGAGCCGCCAGAGCA 2404
Qy 2404 AACAGTCTCTCCCTTGAATTTGAGCAAGAGAGAGTCTGGGATGTGGAAACCTTACT 2463
Db 2405 AACAGTCTCTCTCCCTTGAATTTGAGCAAGAGAGAGTCTGGGATGTGGAAACCTTACT 2464
Qy 2464 CTGTCTCTCTGGAGCAGAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2523
Db 2465 CTGTCTCTCTGGAGCAGAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2524
Qy 2524 GTCCGGAGCAGGAACTGCGAGAGCTGGAATATAGATTTATCTCTCAACAGCCTGTCCAG 2583
Db 2525 GTCCGGAGCAGGAACTGCGAGAGCTGGAATATAGATTTATCTCTCAACAGCCTGTCCAG 2584
Qy 2584 CCATTTCTCTGGAGGAGCAGAGCAATTTCTCTGTCCTCAGCATCGACAGCCTCTCC 2643
Db 2585 CCATTTCTCTGGAGGAGCAGAGCAATTTCTCTGTCCTCAGCATCGACAGCCTCTCC 2644
Qy 2644 CTGTCCGATGACAGTGAGAAGAAACCATCAAGGCTCTCAAGCTCGCGGACACCCCTG 2703
Db 2645 CTGTCCGATGACAGTGAGAAGAAACCATCAAGGCTCTCAAGCTCGCGGACACCCCTG 2704
Qy 2704 AGCTCAGCGTACACTCTCTGGAGAGCAGAGCGGAGGCTCGAAGCTCAGCTGGAACATG 2763
Db 2705 AGCTCAGCGTACACTCTCTGGAGAGCAGAGCGGAGGCTCGAAGCTCAGCTGGAACATG 2764
Qy 2764 GTGCTGGCCCCGGGGCGGCCACAGACACCCCAAGCTATTTCAATGCTGGAAGTCCAA 2823
Db 2765 GTGCTGGCCCCGGGGCGGCCACAGACACCCCAAGCTATTTCAATGCTGGAAGTCCAA 2824
Qy 2824 ATACAGTCTCTTAATGTTGAACACCTGACACATCCGGGAGTTCCACCGGTCGGA 2883
Db 2825 ATACAGTCTCTTAATGTTGAACACCTGACACATCCGGGAGTTCCACCGGTCGGA 2884
Qy 2884 GACATCGCCACTGGCATCAGCAGCAGATCCAGCTGAGGCTTCAAGTGTGTCACCAA 2943
Db 2885 GACATCGCCACTGGCATCAGCAGCAGATCCAGCTGAGGCTTCAAGTGTGTCACCAA 2944
Qy 2944 GACGCGCAGCTGTTCGCTACAGCATGAGGTCAGACTCGGGCATCGACTGCAAGTGC 3003
Db 2945 GACGCGCAGCTGTTCGCTACAGCATGAGGTCAGACTCGGGCATCGACTGCAAGTGC 3004
Qy 3004 ACACCTGGCCCCCTGATGGCAGCTTCCCTTGGAGCTGGAGGGTCAAGCATGGCCAGTGGAG 3063
Db 3005 ACACCTGGCCCCCTGATGGCAGCTTCCCTTGGAGCTGGAGGGTCAAGCATGGCCAGTGGAG 3064
Qy 3064 AACAGGCTTAACCTCTGCTCCACCGCGGCTCCACACTGCGGAAAGAGCCTTCTCTG 3123
Db 3065 AACAGGCTTAACCTCTGCTCCACCGCGGCTCCACACTGCGGAAAGAGCCTTCTCTG 3124
Qy 3124 CTCGCTGACGATGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATTTGCCAGCC 3183
Db 3125 CTCGCTGACGATGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATTTGCCAGCC 3184
Qy 3184 CCCCCTGCTCAGAGTGGGAAACAGGGCTTCGACAGCAAGGTGGGGGCAAGCAAGTGC 3243
Db 3185 CCCCCTGCTCAGAGTGGGAAACAGGGCTTCGACAGCAAGGTGGGGGCAAGCAAGTGC 3244
Qy 3244 TCCAGGATTTCACTTGAAGCTTCCCGCCCTCTCTGCTGAAAGAAATCCCGCCAGTGAA 3303
Db 3245 TCCAGGATTTCACTTGAAGCTTCCCGCCCTCTCTGCTGAAAGAAATCCCGCCAGTGAA 3304
Qy 3304 GAGCAGAGAGGAGTGCAGAGTTACTTGGGGAACAAACACAGGATCTTTTCTGCTCC 3363
Db 3305 GAGCAGAGAGGAGTGCAGAGTTACTTGGGGAACAAACACAGGATCTTTTCTGCTCC 3364

QY 484 GAGTGTGAGAAATAGCCAGAGTTTCAGCCCGACCTTTTCAGAACGATTTTCATCGCTGGG 543
DB 485 GAGTGTGAGAAATAGCCAGAGTTTCAGCCCGACCTTTTCAGAACGATTTTCATCGCTGGG 544
QY 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAATGTGGCCCATGCT 603
DB 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAATGTGGCCCATGCT 604
QY 604 ACAGAGGCGAAATAGCCCGTGTGTGTGGAGGAAAGCGTCGAGCAAAAGCCCGGAAG 663
DB 605 ACAGAGGCGAAATAGCCCGTGTGTGTGGAGGAAAGCGTCGAGCAAAAGCCCGGAAG 664
QY 664 AAACCGAAGAGAGAGTCAAGTCCCTGGTCTATGAGAGTGGCCCTTGCCCAAAACC 723
DB 665 AAACCGAAGAGAGAGTCAAGTCCCTGGTCTATGAGAGTGGCCCTTGCCCAAAACC 724
QY 724 CTCCCCAGAGCCCTGAGCAGAGAGCTGCACATCCAGTCCGAGGAGTGAAGTCTCCA 783
DB 725 CTCCCCAGAGCCCTGAGCAGAGAGCTGCACATCCAGTCCGAGGAGTGAAGTCTCCA 784
QY 784 CTCCGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCTCTGAAAGAACAGGC 843
DB 785 CTCCGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCTCTGAAAGAACAGGC 844
QY 844 CTTCGGCAACTCTGTGTTTAAAGCAGCTTGCGAGAGGCTTACGCCCGCTCTGCCTCGATCA 903
DB 845 CTTCGGCAACTCTGTGTTTAAAGCAGCTTGCGAGAGGCTTACGCCCGCTCTGCCTCGATCA 904
QY 904 GAACTCCACAAACTGATCAGCCCTTGCAATGCTGAAACCAAGTGTGAAACTGCACAC 963
DB 905 GAACTCCACAAACTGATCAGCCCTTGCAATGCTGAAACCAAGTGTGAAACTGCACAC 964
QY 964 CCCCAGGACGAGGCGCCCTGCCCCCTGCCCCAGCCACCTTCCCTTATAGCAGACTGCT 1023
DB 965 CCCCAGGACGAGGCGCCCTGCCCCCTGCCCCAGCCACCTTCCCTTATAGCAGACTGCT 1024
QY 1024 CATCCCTTCCCATTCACCCCTCTCAGCCCTGGAACCTCACCCCTCTGAGTCTTCCTG 1083
DB 1025 CATCCCTTCCCATTCACCCCTCTCAGCCCTGGAACCTCACCCCTCTGAGTCTTCCTG 1084
QY 1084 GGCAAACTGGCTGTGTAGACAGCCAGAAACCTTGGCTTGACCCACACCTGAGCAAACTG 1143
DB 1085 GGCNACTGGCTGTGTAGACAGCCAGAAACCTTGGCTTGACCCACACCTGAGCAAACTG 1144
QY 1144 GCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTTGAGCCAGCTGCCTGTCT 1203
DB 1145 GCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTTGAGCCAGCTGCCTGTCT 1204
QY 1204 CGTGTGCCCATGAGAAAGTTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGCAGC 1263
DB 1205 CGTGTGCCCATGAGAAAGTTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGCAGC 1264
QY 1264 GTGAGCTCAAGCCAGGCGCCACAGCTTACAGCTTGCCCAAGACCTTGGCAGCAGCGGGC 1323
DB 1265 GTGAGCTCAAGCCAGGCGCCACAGCTTACAGCTTGCCCAAGACCTTGGCAGCAGCGGGC 1324
QY 1324 TCCAGATCCCGGAGCGCCAGCCCAAACTGAGGACAAAGAGGTTCTGCTCACTGAG 1383
DB 1325 TCCAGATCCCGGAGCGCCAGCCCAAACTGAGGACAAAGAGGTTCTGCTCACTGAG 1384
QY 1384 AAATCAAGCAGTGGATATGATACCGAAGAGTCCACTTGGCCACGACACAGCTC 1443
DB 1385 AAATCAAGCAGTGGATATGATACCGAAGAGTCCACTTGGCCACGACACAGCTC 1444
QY 1444 CGCTGGCAGAGGCTCTTTCGAGAGGTGCAAGGATGAGGACAAAGCAGACTGGCTTC 1503
DB 1445 CGCTGGCAGAGGCTCTTTCGAGAGGTGCAAGGATGAGGACAAAGCAGACTGGCTTC 1504
QY 1504 CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTTCGGCAGAGAGTGTGAGTGT 1563
DB 1505 CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTTCGGCAGAGAGTGTGAGTGT 1564
QY 1564 GCAGGATTGACCTCACCCAGAAATGTCCCTTTTGTATGGAGCTGTGAGAGAGGGCTTGG 1623

DB 1565 GCAGGATTGACCTCACCCAGAAATGTCCCTTTGTATGGAGCTGTGAGAGAGGGCTTGG 1624
QY 1624 GTCAACATCTTTCATGAGAGTGTGGAAGTGGCTTCCCTGGCCAGCTGGTCAAGAGCAG 1683
DB 1625 GTCAACATCTTTCATGAGAGTGTGGAAGTGGCTTCCCTGGCCAGCTGGTCAAGAGCAG 1684
QY 1684 GGCTGTCTCCACAGAGGACCGGGCCCTGTACTACTCTGGCCAGGCCCTGGAGGGTCTCGAA 1743
DB 1685 GGCTGTCTCCACAGAGGACCGGGCCCTGTACTACTCTGGCCAGGCCCTGGAGGGTCTCGAA 1744
QY 1744 TACCTCCACTCAGAAAGATTCTGCATGGGAGCTCAAAAGCTGACAAAGTGTCTCTGTCC 1803
DB 1745 TACCTCCACTCAGAAAGATTCTGCATGGGAGCTCAAAAGCTGACAAAGTGTCTCTGTCC 1804
QY 1804 AGCAGTGGGAGCAGCAGCCCTCTGTGACTTTGGCCATGCTGTGCTTCAACCTGAT 1863
DB 1805 AGCAGTGGGAGCAGCAGCCCTCTGTGACTTTGGCCATGCTGTGCTTCAACCTGAT 1864
QY 1864 GGCTGGGAAAGTCTTGTCTCACAGGGAGCTACATCTCTGGCACAGAGACCACATGGCT 1923
DB 1865 GGCTGGGAAAGTCTTGTCTCACAGGGAGCTACATCTCTGGCACAGAGACCACATGGCT 1924
QY 1924 CCGAGGTGTGTCTGGGACGAGCTGCGACGCCAAAGTGTGATGTCTGGAGCAGTGTGT 1983
DB 1925 CCGAGGTGTGTCTGGGACGAGCTGCGACGCCAAAGTGTGATGTCTGGAGCAGTGTGT 1984
QY 1984 ATGATGTGCACATGCTCAACGGCTGCGACCCCTGGATCAGTTCCTCCGAGGCGCGCTC 2043
DB 1985 ATGATGTGCACATGCTCAACGGCTGCGACCCCTGGATCAGTTCCTCCGAGGCGCGCTC 2044
QY 2044 TGCCTCAAGATTGCGACGAGCCTCCGCTGTGAGGAGATCCCAACCTCTCTGCGCCCT 2103
DB 2045 TGCCTCAAGATTGCGACGAGCCTCCGCTGTGAGGAGATCCCAACCTCTCTGCGCCCT 2104
QY 2104 CTCACAGCCAGGCCATCCAAAGGGGTGAGGAAAGAGCCCATCCACCGCTGTGTGCA 2163
DB 2105 CTCACAGCCAGGCCATCCAAAGGGGTGAGGAAAGAGCCCATCCACCGCTGTGTGCA 2164
QY 2164 GCGAGCTGGGAGGAGGTGAACCGGGCACTACGCAAGTGGAGGTCTCAAGAGCCCT 2223
DB 2165 GCGAGCTGGGAGGAGGTGAACCGGGCACTACGCAAGTGGAGGTCTCAAGAGCCCT 2224
QY 2224 TGGAGGAGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAACAG 2283
DB 2225 TGGAGGAGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAACAG 2284
QY 2284 ACCCTCCATGCCAGCCGAGAGAGTTTTCGCAAGGGGCCCGAGGCCCGCCAGCTGAG 2343
DB 2285 ACCCTCCATGCCAGCCGAGAGAGTTTTCGCAAGGGGCCCGAGGCCCGCCAGCTGAG 2344
QY 2344 GAGACACAGGACGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAGAGCCA 2403
DB 2345 GAGACACAGGAGGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAGAGCCA 2404
QY 2404 AACAGTCTCTCTCTGACTTTGAGCAAGGAGAGTGTGGGATGTGGGAAACCTTACCT 2463
DB 2405 AACAGTCTCTCTCTGACTTTGAGCAAGGAGAGTGTGGGATGTGGGAAACCTTACCT 2464
QY 2464 CTGTCTCTCTGGAGCAGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2523
DB 2465 CTGTCTCTCTGGAGCAGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2524
QY 2524 GTCCCGGAGCAGGAACTCGACAGCTGGAAATAGAAATATTCTCAACAGCTGTCCAG 2583
DB 2525 GTCCCGGAGCAGGAACTCGACAGCTGGAAATAGAAATATTCTCAACAGCTGTCCAG 2584
QY 2584 CCAATTTCTCTGGAGGAGCAGGAAATTTCTCTGCTCTCAGCATCGACAGCTCTCC 2643
DB 2585 CCAATTTCTCTGGAGGAGCAGGAAATTTCTCTGCTCTCAGCATCGACAGCTCTCC 2644
QY 2644 CTGTGCGATGACAGTGGAGAGAACCCATCAAGGCTCTCAAGCTCCGGGACACCTTC 2703

Db	2645	CTGTCGGATGACAGTGTGAAGAACCCATCAAAAGGCCTCTCAAAGCTCGCGGACACCCCTG	2704
Qy	2704	AGCTCAGCGGTACACTCTCTGGAGCAGCCAGGCCGAGGCTCGAAGCTCCAGCTGCAACATG	2763
Db	2705	AGCTCAGCGGTACACTCTCTGGAGCAGCCAGGCCGAGGCTCGAAGCTCCAGCTGCAACATG	2764
Qy	2764	GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGTTGTAAGTCCAA	2823
Db	2765	GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGTTGTAAGTCCAA	2824
Qy	2824	ATACAGTCTCTTAATGTTGAACACTGCACATCCGGGAGTTCCACCGGTCAAAGTGGGA	2883
Db	2825	ATACAGTCTCTTAATGTTGAACACTGCACATCCGGGAGTTCCACCGGTCAAAGTGGGA	2884
Qy	2884	GACATCGCACTGGCATCAGCAGCAGACATCCAGCTGCAGGCTTCAGCTTGGTCAACAA	2943
Db	2885	GACATCGCACTGGCATCAGCAGCAGACATCCAGCTGCAGGCTTCAGCTTGGTCAACAA	2944
Qy	2944	GACGGGCAGCTGTTTCGCTACGACATGGAGTGCACACTCGGGCATCGACCTCGAGTGC	3003
Db	2945	GACGGGCAGCTGTTTCGCTACGACATGGAGTGCACACTCGGGCATCGACCTCGAGTGC	3004
Qy	3004	ACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAGCATGCCAGCTGGAG	3063
Db	3005	ACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAGCATGCCAGCTGGAG	3064
Qy	3064	AACAGGCCCTAACCTCGCCCTCCACCGCGGCTCCACACTGCCGGAAAGCAGACCTTCCTG	3123
Db	3065	AACAGGCCCTAACCTCGCCCTCCACCGCGGCTCCACACTGCCGGAAAGCAGACCTTCCTG	3124
Qy	3124	CTCGGTGCACGATGCTGCTCCGCTGAAAAACAAGGCTCAGCCGCTCCAGGGGATTCGCCAGCC	3183
Db	3125	CTCGGTGCACGATGCTGCTCCGCTGAAAAACAAGGCTCAGCCGCTCCAGGGGATTCGCCAGCC	3184
Qy	3184	CCC CGGCTCAGCTGGGAAACAGGGCCCTCGCAGCAGCAGAGTGGGGGCACAGCAGATGCC	3243
Db	3185	CCC CGGCTCAGCTGGGAAACAGGGCCCTCGCAGCAGCAGAGTGGGGGCACAGCAGATGCC	3244
Qy	3244	TCCAGGATTTTCACTGTAGCCCTGCCCCACCTGCTGAAAAAAACATCCGCCACGTTAA	3303
Db	3245	TCCAGGATTTTCACTGTAGCCCTGCCCHACCTGCTGAAAAAAACATCCGCCACGTTAA	3304
Qy	3304	GAGACAGAAGGAGGATGGCAGGATTTACTTGGGGAAACAAACAGGGATCTTTTTCTGCC	3363
Db	3305	GAGACAGAAGGAGGATGGCAGGATTTACTTGGGGAAACAAACAGGGATCTTTTTCTGCC	3364
Qy	3364	CCTGCTCAGTGTGAGTTGGCCCTGACACCGCTTGGATCAGTACCAATTTGTTGGCAGACAGG	3423
Db	3365	CCTGCTCAGTGTGAGTTGGCCCTGACACCGCTTGGATCAGTACCAATTTGTTGGCAGACAGG	3424
Qy	3424	GGAGCAGCTTCCAGCTGGTTCAGAAAGGGGTGGCGAGCCCTTCGGCCCTCCACCCCTC	3483
Db	3425	GGAGCAGCTTCCAGCTGGTTCAGAAAGGGGTGGCGAGCCCTTCGGCCCTCCACCCCTC	3484
Qy	3484	CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCCAAATCTCAGGTTTCAAGTGCAGAACCA	3543
Db	3485	CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCCAAATCTCAGGTTTCAAGTGCAGAACCA	3544
Qy	3544	GGTCAGCAGGTATGCCCGCCCGTAAAGTTAAAGGGGGCCCTCTAAACCCCTTCCTGGCCCTC	3603
Db	3545	GGTCAGCAGGTATGCCCGCCCGTAAAGTTAAAGGGGGCCCTCTAAACCCCTTCCTGGCCCTC	3604
Qy	3604	ACCTGGCCAGCTCACCCCTTTTGGTGTGAGGGAAAGAAATGCTGACCTCGGGGAGGCT	3663
Db	3605	ACCTGGCCAGCTCACCCCTTTTGGTGTGAGGGAAAGAAATGCTGACCTCGGGGAGGCT	3664
Qy	3664	CCCTGGTAGAATACACCACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCT	3723
Db	3665	CCCTGGTAGAATACACCACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCT	3724
Qy	3724	GGTTACGCCAAGGACCAAGAGGTGTGTAAAGTGAAGTGTCTTCAAGTCCCCAGACATGT	3783
Db	3725	GGTTACGCCAAGGACCAAGAGGTGTGTAAAGTGAAGTGTCTTCAAGTCCCCAGACATGT	3784

[illegible]

; CURRENT FILING DATE: 2003-01-15									
; PRIOR APPLICATION NUMBER: 60/298,918									
; PRIOR FILING DATE: 2001-06-18									
; PRIOR APPLICATION NUMBER: 60/380,710									
; PRIOR FILING DATE: 2002-05-14									
; PRIOR APPLICATION NUMBER: 10/172,118									
; PRIOR FILING DATE: 2002-06-14									
; NUMBER OF SEQ ID NOS: 2699									
; SEQ ID NO 975									
; LENGTH: 4596									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-10-342-887-975									
Query Match									
Best Local Similarity 99.6%; Score 4578.4; DB 52; Length 4596;									
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;									
QY	5	GGGGACTGTGCGGTGTGGAAACGTGTAGCTGTGTGA-AGGTGACTCTCTTTACCAATTGAGG	63						
Db	5	GGGGACTGTGCGGTGTGGAAACGTGTAGCTGTGTGAGAGGTGACTCTCTTTACCAATTGAGG	64						
QY	64	ATGTTTGGAGATGAGTATGTGTGCGAGGACACATATAACAGCAGACAGACCTTTGCC	123						
Db	65	ATGTTTGGAGATGAGTATGTGTGCGAGGACACATATAACAGCAGACAGACCTTTGCC	124						
QY	124	CTGCGCTTTCTCCCCCAACCCAGGCTGACCTGTGTCTCTCCAGGTCTGGGATTTCTAAGT	183						
Db	125	CTGCGCTTTCTCCCCCAACCCAGGCTGACCTGTGTCTCTCCAGGTCTGGGATTTCTAAGT	184						
QY	184	GACCTGCTCTGTGTTTGTGTTCTCTCAGATGAGCACAAAGCTCTGGAGATGGCAGTGATG	243						
Db	185	GACCTGCTCTGTGTTTGTGTTCTCTCAGATGAGCACAAAGCTCTGGAGATGGCAGTGATG	244						
QY	244	GAAATGGCTGCGCAGGTGCCCCCTGGCTCAGCAGTGGGGCAGCAGAAAGTCCCGAAG	303						
Db	245	GAAATGGCTGCGCAGGTGCCCCCTGGCTCAGCAGTGGGGCAGCAGAAAGTCCCGAAG	304						
QY	304	CCAAAGGAGAAAGACCGCCGCACTGGGGAAGAAACAGAGCTCCGCTCTCAAGCTTGAGGCC	363						
Db	305	CCAAAGGAGAAAGACCGCCGCACTGGGGAAGAAACAGAGCTCCGCTCTCAAGCTTGAGGCC	364						
QY	364	GTGAGAAAGAGCCCTGTGTTCTCGGAAAGTGGAGATCCTGAATGACGTGATTAACCAAG	423						
Db	365	GTGAGAAAGAGCCCTGTGTTCTCGGAAAGTGGAGATCCTGAATGACGTGATTAACCAAG	424						
QY	424	GGCAGCAAGAGAGGTCCGAGCGAGGCCAGCTGCCATCTCTATCATCGCCAGGCT	483						
Db	425	GGCAGCAAGAGAGGTCCGAGCGAGGCCAGCTGCCATCTCTATCATCGCCAGGCT	484						
QY	484	GAGTGTGAGAAATAGCCAGAGTTCAAGCCCACTTTTCAAGACGCAATTTTCATCGCTGGG	543						
Db	485	GAGTGTGAGAAATAGCCAGAGTTCAAGCCCACTTTTCAAGACGCAATTTTCATCGCTGGG	544						
QY	544	TCCAAACAGTACAGCAGGTCCGAGAGTCTGTATCATAGATCCCAACAAATGTGGCCCATGCT	603						
Db	545	TCCAAACAGTACAGCAGGTCCGAGAGTCTGTATCATAGATCCCAACAAATGTGGCCCATGCT	604						
QY	604	ACAGAGGGCAAAATGGCCCGTGTGTTGGAAAGGAAAGCGTCGACGAAAGCCCGGAAG	663						
Db	605	ACAGAGGGCAAAATGGCCCGTGTGTTGGAAAGGAAAGCGTCGACGAAAGCCCGGAAG	664						
QY	664	AAAACGGAAGAAAGAGTCAAGTCCCTGGCTCATGACAGAGTGGCCCTTGGCCAAACCC	723						
Db	665	AAAACGGAAGAAAGAGTCAAGTCCCTGGCTCATGACAGAGTGGCCCTTGGCCAAACCC	724						
QY	724	CTCCCCAGACCCCTGAGCAGAGAGCTGCACCATCCCAAGTCCAGAGGATGAGTCTCCA	783						
Db	725	CTCCCCAGACCCCTGAGCAGAGAGCTGCACCATCCCAAGTGCAGAGGATGAGTCTCCA	784						
QY	784	CTCGGGCCCAATATGTTAGAAACACCCCGCAGTTCACCAAGCCTCTGAAAGAAACAGGC	843						
Db	785	CTCGGGCCCAATATGTTAGAAACACCCCGCAGTTCACCAAGCCTCTGAAAGAAACAGGC	844						

QY	844	CTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGGCCTTACGGCCGGCTCTGCTCTGATCA	903						
Db	845	CTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGGCCTTACGGCCGGCTCTGCTCTGATCA	904						
QY	904	GAACTCCACAAACTGATCAGCCCTTGCATATGTTCTGAACCAAGCTGTGAAACTGCACCAC	963						
Db	905	GAACTCCACAAACTGATCAGCCCTTGCATATGTTCTGAACCAAGCTGTGAAACTGCACCAC	964						
QY	964	CCCAGGACGAGGCCCCCTGCTGCCACGACCCCTTCCCTTCCCTATAGCAGACTGCT	1023						
Db	965	CCCAGGACGAGGCCCCCTGCTGCCACGACCCCTTCCCTTCCCTATAGCAGACTGCT	1024						
QY	1024	CATCCCTTCCCAATTCACCCCTCTCAGCCCTTGGAAACCTCAACCTCTCGAGTCTTCTG	1083						
Db	1025	CATCCCTTCCCAATTCACCCCTCTCAGCCCTTGGAAACCTCAACCTCTCGAGTCTTCTG	1084						
QY	1084	GGCAAACTGGGCTGTGTAGACAGCCAGCAAAACCTTCCCTGACCCACACCTGAGCAACTG	1143						
Db	1085	GGCAAACTGGGCTGTGTAGACAGCCAGCAAAACCTTCCCTGACCCACACCTGAGCAAACTG	1144						
QY	1144	GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGCT	1203						
Db	1145	GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGCT	1204						
QY	1204	CGTGTGCTCCATGAGAAAGTCTTGTGTGAGGAATACTAGTGCAATGCTCTGCAAGGCAGC	1263						
Db	1205	CGTGTGCTCCATGAGAAAGTCTTGTGTGAGGAATACTAGTGCAATGCTCTGCAAGGCAGC	1264						
QY	1264	GTGAGCTCAAGCCAGGCCCCACAGCTGACCCAGCTGGCCAAAGCTTGGGACGACGGGGC	1323						
Db	1265	GTGAGCTCAAGCCAGGCCCCACAGCTGACCCAGCTGGCCAAAGCTTGGGACGACGGGGC	1324						
QY	1324	TCCAGATCCCGGAGCCAGCCCAAAACTGAGGACAAAGGGGTCTCTGCTCACTGAG	1383						
Db	1325	TCCAGATCCCGGAGCCAGCCCAAAACTGAGGACAAAGGGGTCTCTGCTCACTGAG	1384						
QY	1384	AAACTCAAGCCAGTGGATTATGAGTACCGAGAAGAAAGTCCACTGGGCCACCGCACAGCTC	1443						
Db	1385	AAACTCAAGCCAGTGGATTATGAGTACCGAGAAGAAAGTCCACTGGGCCACCGCACAGCTC	1444						
QY	1444	CGCTGGGACAGAGCTCTCTCGAGAGGTGACAGGATGGAGGACAAAGCAGACTGGCTTC	1503						
Db	1445	CGCTGGGACAGAGCTCTCTCGAGAGGTGACAGGATGGAGGACAAAGCAGACTGGCTTC	1504						
QY	1504	CAGTGCCTGTCAAAGAGGTGCGCTGGAAGTATTTTCGGGACAGAGAGCTGATGGCATGT	1563						
Db	1505	CAGTGCCTGTCAAAGAGGTGCGCTGGAAGTATTTTCGGGACAGAGAGCTGATGGCATGT	1564						
QY	1564	GCAGGATTGACCTCAACCCAGAAATTTGCTCCCTTTGATGGAGCTGTGAGAAAGGGCTTGG	1623						
Db	1565	GCAGGATTGACCTCAACCCAGAAATTTGCTCCCTTTGATGGAGCTGTGAGAAAGGGCTTGG	1624						
QY	1624	GTCAACATCTTCAAGAGCTGTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGAGCAG	1683						
Db	1625	GTCAACATCTTCAAGAGCTGTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGAGCAG	1684						
QY	1684	GGCTGTCTCCAGAGGACCGGCGCTGTACTACCTGGGCGAGGCTTGGAGGTCTGGA	1743						
Db	1685	GGCTGTCTCCAGAGGACCGGCGCTGTACTACCTGGGCGAGGCTTGGAGGTCTGGA	1744						
QY	1744	TACCTCCACTCAAGAGGATTTGATGGGGAAGTCAAGCTGACAAAGCTGCAAGCTGCTGTC	1803						
Db	1745	TACCTCCACTCAAGAGGATTTGATGGGGAAGTCAAGCTGACAAAGCTGCAAGCTGCTGTC	1804						
QY	1804	AGCGATGGAGCCAGCGACCTCTGTGACTTTGGCCATGCTGTGTGCTTCAACTGAT	1863						
Db	1805	AGCGATGGAGCCAGCGACCTCTGTGACTTTGGCCATGCTGTGTGCTTCAACTGAT	1864						
QY	1864	GGCTGGGAAAGTCTTGTCTCAGAGGACTTACATCTTGGGACAGAGACCCACATGGCT	1923						
Db	1865	GGCTGGGAAAGTCTTGTCTCAGAGGACTTACATCTTGGGACAGAGACCCACATGGCT	1924						

QY 1924 CCGAGGTGGTGGTGGGAGAGCTGGACGCCAAGGTGGATGTCTGGAGCAGCTGCTGT 1983
Db 1925 CCGAGGTGGTGGTGGGAGAGCTGGACGCCAAGGTGGATGTCTGGAGCAGCTGCTGT 1984
QY 1984 ATGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAAGTTCTTCGAGGCGCGCTC 2043
Db 1985 ATGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAAGTTCTTCGAGGCGCGCTC 2044
QY 2044 TGCCTCAAGATTGCCAGGAGCTCCGCTCTGTGAGGAGATCCACCCCTCTCTGCGCCCT 2103
Db 2045 TGCCTCAAGATTGCCAGGAGCTCCGCTCTGTGAGGAGATCCACCCCTCTCTGCGCCCT 2104
QY 2104 CTCACAGCCAGGCCATCTCAAGAGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2163
Db- 2105 CTCACAGCCAGGCCATCTCAAGAGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2164
QY 2164 GCGAGCTGGAGGGAAGGTGAACCGGCACTACAGCAAGTGGAGGTCTGAAGAGCCCT 2223
Db 2165 GCGAGCTGGAGGGAAGGTGAACCGGCACTACAGCAAGTGGAGGTCTGAAGAGCCCT 2224
QY 2224 TGGAGGGGAGGAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTAACACAG 2283
Db 2225 TGGAGGGGAGGAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTAACACAG 2284
QY 2284 ACCCTCCATGCCAGCGAGAGCTTTGGCCAAAGGGCCCAAGGGCCCGGCCACAGCTGAG 2343
Db 2285 ACCCTCCATGCCAGCGAGAGCTTTGGCCAAAGGGCCCAAGGGCCCGGCCACAGCTGAG 2344
QY 2344 GAGACACAGCAGAGCCCTTAAGCTCCAGCTCCCTCCACAGAGGGCCCAAGGGCCCGGCCACAGAGCCA 2403
Db 2345 GAGACACAGCAGAGCCCTTAAGCTCCAGCTCCCTCCACAGAGGGCCCAAGGGCCCGGCCACAGAGCCA 2404
QY 2404 AACAAAGTCTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCTTACCT 2463
Db 2405 AACAAAGTCTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCTTACCT 2464
QY 2464 CTGTCCTCTCTGGAGCAGAGCCCTTGCAGAAACCCAGCTCAACAGAGGGCCCAAGAGCAAC 2523
Db 2465 CTGTCCTCTCTGGAGCAGAGCCCTTGCAGAAACCCAGCTCAACAGAGGGCCCAAGAGCAAC 2524
QY 2524 GTCCCGGAGGAGGAGTGCAGCAGCTGGAATAGAAATTTCTCAACAGAGCTTCCAG 2583
Db 2525 GTCCCGGAGGAGGAGTGCAGCAGCTGGAATAGAAATTTCTCAACAGAGCTTCCAG 2584
QY 2584 CCATTTTCTCTGGAGGAGCAGAGCAATTTCTCTGCTGCTCAGCATCTGAGCTCTCC 2643
Db 2585 CCATTTTCTCTGGAGGAGCAGAGCAATTTCTCTGCTGCTCAGCATCTGAGCTCTCC 2644
QY 2644 CTGTCGGATGACAGTGAGAGAACCCATCAAGGCCCTCTCAAGCTCGCGGACACCCCTG 2703
Db 2645 CTGTCGGATGACAGTGAGAGAACCCATCAAGGCCCTCTCAAGCTCGCGGACACCCCTG 2704
QY 2704 AGCTCAGGCTGACTCTCTGGAGCAGAGCGCCAGAGCTGGAAGCTCAGCTGGAACATG 2763
Db 2705 AGCTCAGGCTGACTCTCTGGAGCAGAGCGCCAGAGCTGGAAGCTCAGCTGGAACATG 2764
QY 2764 GTGCTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAGTCCAA 2823
Db 2765 GTGCTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAGTCCAA 2824
QY 2824 ATA CAGTCTCTTAATGTGTGAACACCTGCACATCCGGGAGTTTCAACCGGGTCAAGAGTGGGA 2883
Db 2825 ATA CAGTCTCTTAATGTGTGAACACCTGCACATCCGGGAGTTTCAACCGGGTCAAGAGTGGGA 2884
QY 2884 GACATCGGCATCGGCATCAGCAGCCAGATCCAGCTGAGGCTTTCAGCTTGGTCAACAAA 2943
Db 2885 GACATCGGCATCGGCATCAGCAGCCAGATCCAGCTGAGGCTTTCAGCTTGGTCAACAAA 2944
QY 2944 GACGGGAGCGCTGTTGCTACGACATGAGGTGCCAGCTCGGGCATCGACCTGCACTGC 3003
Db 2945 GACGGGAGCGCTGTTGCTACGACATGAGGTGCCAGCTCGGGCATCGACCTGCACTGC 3004
QY 3004 AACTGGCCCCCTGATGGCAGCTTCCCTGGAGCTGGAGGGGTCAAGCATGGCCAGCTGGAG 3063

Db 3005 AACTGGCCCCCTGATGGCAGCTTCCCTGGAGCTGAGGGGTCAAGCATGGCCAGCTGGAG 3064
QY 3064 AACAGGCCCCATAACCTCCCTCCACCGCGGCTCCACACTGCGCGGAAAGAGCCTTCTCTG 3123
Db 3065 AACAGGCCCCATAACCTCCCTCCACCGCGGCTCCACACTGCGCGGAAAGAGCCTTCTCTG 3124
QY 3124 CTGGTGCACGATGCTGCCCTGAAAAACACAGGCTCAGCGCTTCCACAGGGGATGCGCAGCC 3183
Db 3125 CTGGTGCACGATGCTGCCCTGAAAAACACAGGCTCAGCGCTTCCACAGGGGATGCGCAGCC 3184
QY 3184 CCGCGCTCAGCTGGGAAACAGGGCTCCAGCAGCAAGGTGGGGGCAAGCAGAAATGCC 3243
Db 3185 CCGCGCTCAGCTGGGAAACAGGGCTCCAGCAGCAAGGTGGGGGCAAGCAGAAATGCC 3244
QY 3244 TCCCAGGATTTTCAACCTGAGCCCTGCCACCTGCTGAAAAAACAATCCGCCACGTGAA 3303
Db 3245 TCCCAGGATTTTCAACCTGAGCCCTGCCACCTGCTGAAAAAACAATCCGCCACGTGAA 3304
QY 3304 GAGACAGAGGAGATGGCAGGAGTTTACCTGGGAAACAAACAGGAGTCTTTTCTGCC 3363
Db 3305 GAGACAGAGGAGATGGCAGGAGTTTACCTGGGAAACAAACAGGAGTCTTTTCTGCC 3364
QY 3364 CTTGCTCAGTCCAGTTGGCTGACCCGCTTGGATCAGTGACCAATTTTGGGCAAGAG 3423
Db 3365 CTTGCTCAGTCCAGTTGGCTGACCCGCTTGGATCAGTGACCAATTTTGGGCAAGAG 3424
QY 3424 GGAGAGCAGCTTCCAGCTGGGTGAGAGGGGTGGGCGAGCCCTTCCGCCCTCACCCCTC 3483
Db 3425 GGAGAGCAGCTTCCAGCTGGGTGAGAGGGGTGGGCGAGCCCTTCCGCCCTCACCCCTC 3484
QY 3484 CAGGCTCTCTGAGAGTGTCAAGTGTAAAGGGGCCAAACTCAGGTTTCAAGTGCAGAACCA 3543
Db 3485 CAGGCTCTCTGAGAGTGTCAAGTGTAAAGGGGCCAAACTCAGGTTTCAAGTGCAGAACCA 3544
QY 3544 GGTTCAGCAGTATGCCCGCCCTTAGGTTAAGGGGCCCTCTAAACCCCTTCCCTGGCCCTC 3603
Db 3545 GGTTCAGCAGTATGCCCGCCCTTAGGTTAAGGGGCCCTCTAAACCCCTTCCCTGGCCCTC 3604
QY 3604 ACCTGGCCAGCTCACCCCTTTGGGTGAGGGGAAAGAAATGCCCTGACCCCTGGGAAGGCT 3663
Db 3605 ACCTGGCCAGCTCACCCCTTTGGGTGAGGGGAAAGAAATGCCCTGACCCCTGGGAAGGCT 3664
QY 3664 CCCTGTGAGTAATACACACTTTTTCAGGTTTGTGAACACACAGGCTCTGAGTTGACCTCT 3723
Db 3665 CCCTGTGAGTAATACACACTTTTTCAGGTTTGTGAACACACAGGCTCTGAGTTGACCTCT 3724
QY 3724 GGTTCAGCCAAAGGACCAAGAGGTGTAAAGTGAAGTGGTTCTCAGTCCCGACACATGT 3783
Db 3725 GGTTCAGCCAAAGGACCAAGAGGTGTAAAGTGAAGTGGTTCTCAGTCCCGACACATGT 3784
QY 3784 GCGCTTTGCTGTGGCTACCACTTCCAGAGCAGAGGCCCGCCAGCCCTTTCAGGC 3843
Db 3785 GCGCTTTGCTGTGGCTACCACTTCCAGAGCAGAGGCCCGCCAGCCCTTTCAGGC 3844
QY 3844 CCAGCACTGCCCCAGACTCCGCTGCACTCAGTTTCCCTCATCTGTAAAGGTGAAGGGTGTAT 3903
Db 3845 CCAGCACTGCCCCAGACTCCGCTGCACTCAGTTTCCCTCATCTGTAAAGGTGAAGGGTGTAT 3904
QY 3904 GCAGGATATCCCTGACAGGAACAGTCTGTGGATGGACATGATCAGTCTTAAGGAAAGCAG 3963
Db 3905 GCAGGATATCCCTGACAGGAACAGTCTGTGGATGGACATGATCAGTCTTAAGGAAAGCAG 3964
QY 3964 CAGAGAGAGCTCCCGGCCCGCCAGCCCACTATCAGTGTCCAGCGTGTGTTCCCCAG 4023
Db 3965 CAGAGAGAGCTCCCGGCCCGCCAGCCCACTATCAGTGTCCAGCGTGTGTTCCCCAG 4024
QY 4024 AGCAGACTCAGCATCACACTGACACTCACCTGCGCCCTGCCCGCCAGAGGCTACTGC 4083
Db 4025 AGCAGACTCAGCATCACACTGACACTCACCTGCGCCCTGCCCGCCAGAGGCTACTGC 4084
QY 4084 CGACGGCACTTTCATCTGATGACCTCAAGCACTTTTCATGCTGCTGCTGCGAGGGC 4143

Db 4085 CGACGGCACTTTGCACTCTGATGACCTCAAGACATTTTCATGGCTGCCCCCTCTGGCAGGCG 4144
QY 4144 AGGGCAGGCGAGTGAACACTGTAGGACATAGCAAGCCAGAGATGGGGTGAAGGACACA 4203
Db 4145 AGGGCAGGCGAGTGAACACTGTAGGACATAGCAAGCCAGAGATGGGGTGAAGGACACA 4204
QY 4204 GTCTTGAGCTGTCCACATGCATGTGACTCCTCAAACCTCTTCCAGATTCTCTAAGAATA 4263
Db 4205 GTCTTGAGCTGTCCACATGCATGTGACTCCTCAAACCTCTTCCAGATTCTCTAAGAATA 4264
QY 4264 GCACCCCTTCCCATTCGCCAGCTTAGCTCTTCTCCAGGGAGAGCTACTCAGACTC 4323
Db 4265 GCACCCCTTCCCATTCGCCAGCTTAGCTCTTCTCCAGGGAGAGCTACTCAGACTC 4324
QY 4324 ACGTAGCATTAATCAGCTGTGAATCGTCAGGGGTGTCTGTAGCTCAACCTCTCTGG 4383
Db 4325 ACGTAGCATTAATCAGCTGTGAATCGTCAGGGGTGTCTGTAGCTCAACCTCTCTGG 4384
QY 4384 GCAGGGAGCGCGAGACTCCGTGGGAGAGCTCATTTCCACATCTTGCCAAGACAGCTT 4443
Db 4385 GCAGGGAGCGCGAGACTCCGTGGGAGAGCTCATTTCCACATCTTGCCAAGACAGCTT 4444
QY 4444 TGTCCAGCTGTCCACATTCAGTTCAGCTCAGCTCTCCGGGGAGAGAGCCCGGCCCCAGCAC 4503
Db 4445 TGTCCAGCTGTCCACATTCAGTTCAGCTCAGCTCTCCGGGGAGAGAGCCCGGCCCCAGCAC 4504
QY 4504 ATAAAGAACTGCAGCTTGTGTAATCGTCAGGGGTGTCTGTAGCTCAACCTCTCTGG 4563
Db 4505 ATAAAGAACTGCAGCTTGTGTAATCGTCAGGGGTGTCTGTAGCTCAACCTCTCTGG 4564
QY 4564 TAAAGTTTGGGTGATGACAAATGTTTAAAAA 4595
Db 4565 TAAAGTTTGGGTGATGACAAATGTTTAAAAA 4596

RESULT 12

US-60-474-850-144
; Sequence 144, Application US/60474850
; GENERAL INFORMATION:
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING
; FILE REFERENCE: 506613000700
; CURRENT APPLICATION NUMBER: US/60/474,850
; CURRENT FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 552
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 144
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-474-850-144

Query Match 99.6%; Score 4578.4; DB 117; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 5 GGGGACCTGTGCGGTGTGAACAGTGTAGCTGTGTGA-AGTGGACTCTGTGTACCATTGAGG 63
Db 5 GGGGACCTGTGCGGTGTGAACAGTGTAGCTGTGTGAAGGTGGACTCTGTGTACCATTGAGG 64
QY 64 ATGTTTGGAGGATGATGATGTGTCAGAGGACACATAAACAGGACAGACCTTTGCC 123
Db 65 ATGTTTGGAGGATGATGATGTGTCAGAGGACACATAAACAGGACAGACCTTTGCC 124
QY 124 CCTGCTTTCTCCCCAACCCCAAGGCTGACCTGTGTTCTCCAGGCTCTGGGATTTCTAAGT 183
Db 125 CCTGCTTTCTCCCCAACCCCAAGGCTGACCTGTGTTCTCCAGGCTCTGGGATTTCTAAGT 184
QY 184 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGACACAGCTGGGAGATGGCAGTGATG 243

Db 185 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAAGCTGGGAGATGGCAGTGATG 244
QY 244 GAAATGGGCTGCCAGGTGCCCTCTGGCTCAGCAGTGGGGCAGCAGAAAGAACTCCCCAG 303
Db 245 GAAATGGGCTGCCAGGTGCCCTCTGGCTCAGCAGTGGGGCAGCAGAAAGAACTCCCCAG 304
QY 304 CCAAAAGGAGAGACCCGCCCACTGGGGAGAAACAGAGCTCCGTCTCAAAAGCTTGAGGCC 363
Db 305 CCAAAAGGAGAGACCCGCCCACTGGGGAGAAACAGAGCTCCGTCTCAAAAGCTTGAGGCC 364
QY 364 GTGGAGAAAGAGCCCTGTGTTCTCGGAAAGTGGGAGATCTTGAATGAGCTGATTAACAG 423
Db 365 GTGGAGAAAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCTTGAATGAGCTGATTAACAG 424
QY 424 GGCACAGCCAAAGGAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATCATCGCCAGGCT 483
Db 425 GGCACAGCCAAAGGAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATCATCGCCAGGCT 484
QY 484 GAGTGTGAGAAATAGCCAAAGATTCAAGCCCACTTTTTCAGAAACGCAATTTTCATCGCTGG 543
Db 485 GAGTGTGAGAAATAGCCAAAGATTCAAGCCCACTTTTTCAGAAACGCAATTTTCATCGCTGG 544
QY 544 TCCAAAAGTACAGCCAGCTCCGAGAGTCTTGAATCAGATCCCCCAACAAATGTGGCCCATGCT 603
Db 545 TCCAAAAGTACAGCCAGCTCCGAGAGTCTTGAATCAGATCCCCCAACAAATGTGGCCCATGCT 604
QY 604 ACAGAGGGCAAAATGCGCCGCTGTGTTGGAGGGAGAAAGCGTCGAGCAAAAGCCCGGAAG 663
Db 605 ACAGAGGGCAAAATGCGCCGCTGTGTTGGAGGGAGAAAGCGTCGAGCAAAAGCCCGGAAG 664
QY 664 AAAACGGAAGAAAGAGAGCTCAAAGTCCCTGCTCTATGACAGGATGGCTTGGCCAAACCC 723
Db 665 AAAACGGAAGAAAGAGAGCTCAAAGTCCCTGCTCTATGACAGGATGGCTTGGCCAAACCC 724
QY 724 CTCCCGCAGAGACCCCTGTAGCAGAGAGCTGACACATCCCAAGTGCAGGAGATGAGTCTCA 783
Db 725 CTCCCGCAGAGACCCCTGTAGCAGAGAGCTGACACATCCCAAGTGCAGGAGATGAGTCTCA 784
QY 784 CTGGCGGCCCATATGTTAGAAACACACCGCAGTTTACAAAGCCCTCTCAAGAAACAGGC 843
Db 785 CTGGCGGCCCATATGTTAGAAACACACCGCAGTTTACAAAGCCCTCTCAAGAAACAGGC 844
QY 844 CTGGCGCAACTCTGTTTTAAGCAGCTTGGCAGGAGCCCTACGGCCGGCTCTGCTCGATCA 903
Db 845 CTGGCGCAACTCTGTTTTAAGCAGCTTGGCAGGAGCCCTACGCCGGCTCTGCTCGATCA 904
QY 904 GAACTCCAAACAGTATCAGCCCTTGCATGTTCTGAACCAAGTGTGAAACTGCAACAC 963
Db 905 GAACTCCAAACAGTATCAGCCCTTGCATGTTCTGAACCAAGTGTGAAACTGCAACAC 964
QY 964 CCCCAGGACGGAGGCCCTGCGCCCTGCGCCACGACCCCTTCCCTTATAGCAGACTGCT 1023
Db 965 CCCCAGGACGGAGGCCCTGCGCCCTGCGCCACGACCCCTTCCCTTATAGCAGACTGCT 1024
QY 1024 CATCCCTTCCATTCACCCCTCTCAGCCCTTGGAAACCTCAACCTCTCGAGTCTTCTG 1083
Db 1025 CATCCCTTCCATTCACCCCTCTCAGCCCTTGGAAACCTCAACCTCTCGAGTCTTCTG 1084
QY 1084 GGCAAAATGGGCTGTGTAGACAGCAGAAAACCTTGGCTGTGACCCACACACTGAGCAACTG 1143
Db 1085 GGCAAAATGGGCTGTGTAGACAGCAGAAAACCTTGGCTGTGACCCCACTCTGAGCAACTG 1144
QY 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTTGGAGCCAGCTGCTGCT 1203
Db 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTTGGAGCCAGCTGCTGCT 1204
QY 1204 CGTGTGCCCATGAGAAAGTCTTGTGTGAGGAATACTTAGTGCATGCTCTGCAAGGACG 1263
Db 1205 CGTGTGCCCATGAGAAAGTCTTGTGTGAGGAATACTTAGTGCATGCTCTGCAAGGACG 1264
QY 1264 GTGAGCTCAAGCCAGGCCCAAGCTGACAGCTGGCCAGACTGGGACAGCTGGGACGAGG 1323
Db 1265 GTGAGCTCAAGCCAGGCCCAAGCTGACAGCTGGCCCAAGACTGGGACAGCTGGGACGAGG 1324

1324 TCCAGATCCCGGAGCCAGCCCCCAAACTGAGGACAAACGAGGTGTCTGCTCACTGAG 1383
1325 TCCAGATCCCGGAGCCAGCCCCCAAACTGAGGACAAACGAGGTGTCTGCTCACTGAG 1384
1384 AAACTCAAGCCAGTGTATGATACCGAGAGAGTCCACTGGGCGACGACCAAGTTC 1443
1385 AAACTCAAGCCAGTGTATGATACCGAGAGAGTCCACTGGGCGACGACCAAGTTC 1444
1444 CGCTGGGCGAGGCTCTCTCGAGAGGTGCACAGGATGAGGAGTCCACTGGGCGACGACCAAGTTC 1503
1445 CGCTGGGCGAGGCTCTCTCGAGAGGTGCACAGGATGAGGAGTCCACTGGGCGACGACCAAGTTC 1504
1504 CAGTCCGCTGTCAAAAGGTCCGCTTGAAGTATTTTGGGCGAGGAGTGTATGGCATGT 1563
1505 CAGTCCGCTGTCAAAAGGTCCGCTTGAAGTATTTTGGGCGAGGAGTGTATGGCATGT 1564
1564 GCAGGATGACCTCACCCAGAAATTTGCTCCCTTGTATGAGAGTGTGAGAGAGGCTTGG 1623
1565 GCAGGATGACCTCACCCAGAAATTTGCTCCCTTGTATGAGAGTGTGAGAGAGGCTTGG 1624
1624 GTCAACATCTTCAATGAGTCTCTGGAAGTGTGCTCCCTGGGCGACGACCAAGTTC 1683
1625 GTCAACATCTTCAATGAGTCTCTGGAAGTGTGCTCCCTGGGCGACGACCAAGTTC 1684
1684 GGCTGTCTCCAGAGGACCGGGCCCTGTGTACTCTGGGCGAGGCTTGAAGTTC 1743
1685 GGCTGTCTCCAGAGGACCGGGCCCTGTGTACTCTGGGCGAGGCTTGAAGTTC 1744
1744 TACCTCCATCAGAGAGATTTCTGATGGGAGCTCAAGTGTGACAGGCTCTCTGTCC 1803
1745 TACCTCCATCAGAGAGATTTCTGATGGGAGCTCAAGTGTGACAGGCTCTCTGTCC 1804
1804 AGCGATGGGACCGAGGCTCTGTGACTTTGGCCATGTGTGTCTTCAACCTGAT 1863
1805 AGCGATGGGACCGAGGCTCTGTGACTTTGGCCATGTGTGTCTTCAACCTGAT 1864
1864 GGCTGGGAAAGTCTTCTCTCACAGGAGTCTACATCTCTGGGCGAGGCTTGAAGTTC 1923
1865 GGCTGGGAAAGTCTTCTCTCACAGGAGTCTACATCTCTGGGCGAGGCTTGAAGTTC 1924
1924 CCGAGGTGTGTCTGGGCGAGGCTTCAAGGAGTGTGATGTCTGGAGGAGTGTGTGT 1983
1925 CCGAGGTGTGTCTGGGCGAGGCTTCAAGGAGTGTGATGTCTGGAGGAGTGTGTGT 1984
1984 ATGATGTGTGATGTCTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 2043
1985 ATGATGTGTGATGTCTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 2044
2044 TGCCTCAAGATTTGCGAGGAGCTTCCGCTGTGAGGAGATCCCAACCTCTCTGGGCGCT 2103
2045 TGCCTCAAGATTTGCGAGGAGCTTCCGCTGTGAGGAGATCCCAACCTCTCTGGGCGCT 2104
2104 CTCACAGCCCGAGGCTTCAAGGAGGCTTGAAGGAGGCTTCAAGGAGGCTTGAAGG 2163
2105 CTCACAGCCCGAGGCTTCAAGGAGGCTTGAAGGAGGCTTCAAGGAGGCTTGAAGG 2164
2164 GCGAGTGTGGAGGAGGAGTGAACCGGCTTCAAGGAGGCTTGAAGGAGGCTTGAAGG 2223
2165 GCGAGTGTGGAGGAGGAGTGAACCGGCTTCAAGGAGGCTTGAAGGAGGCTTGAAGG 2224
2224 TGGAGGAGGAGATATAAGAGAACCAAGATCCACCGGCAATCAAGGAGGCTTGAAGG 2283
2225 TGGAGGAGGAGATATAAGAGAACCAAGATCCACCGGCAATCAAGGAGGCTTGAAGG 2284
2284 ACCCTTCCATGCCAGCCGAGAGGCTTTTGGCCAAAGGCTTCCAGGAGGCTTCCAGG 2343
2285 ACCCTTCCATGCCAGCCGAGAGGCTTTTGGCCAAAGGCTTCCAGGAGGCTTCCAGG 2344
2344 GAGACAAAGGAGGAGGCTTGAAGTCCAGGCTTCTCTCCACGAGGCTTCCAGGAGG 2403
2345 GAGACAAAGGAGGAGGCTTGAAGTCCAGGCTTCTCTCCACGAGGCTTCCAGGAGG 2404

2404 AACAGTCTCTCTCCCTTGACTTTGAGCAAGGAGGTCTGGGATGTGGGACCTTACCT 2463
2405 AACAGTCTCTCTCTCCCTTGACTTTTGGCAAGGAGGTCTGGGATGTGGGAACTTACCT 2464
2464 CTGTCTCTCTCTGGAGCAGCCCTGTCAGAAAACCCAGCTTCAACAGAGGAGGAAAGCAACC 2523
2465 CTGTCTCTCTCTGGAGCAGCCCTGTCAGAAAACCCAGCTTCAACAGAGGAGGAAAGCAACC 2524
2524 GTCCGAGCAGGAATCTGAGCAGCTGGAATAGAATTAATTTCTTCAACAGGCTTCCAG 2583
2525 GTCCGAGCAGGAATCTGAGCAGCTGGAATAGAATTAATTTCTTCAACAGGCTTCCAG 2584
2584 CCATTTCTCTGGAGGAGGAGCAAAATTTCTCTGTCCTCAGCATCGACGCTCTCC 2643
2585 CCATTTCTCTGGAGGAGGAGCAAAATTTCTCTGTCCTCAGCATCGACGCTCTCC 2644
2644 CTGTCCGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCGCGGACACCTCTG 2703
2645 CTGTCCGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCGCGGACACCTCTG 2704
2704 AGCTCAGGCTTACATCTCTGGAGCAGGCGAGGCTCGAAGCTCCAGCTGGAACATG 2763
2705 AGCTCAGGCTTACATCTCTGGAGCAGGCGAGGCTCGAAGCTCCAGCTGGAACATG 2764
2764 GTCTGCGCCCGGGCGGCGCCACCGACACCCCAAGCTTATTTCAATGTGTGAAGTCCAA 2823
2765 GTCTGCGCCCGGGCGGCGCCACCGACACCCCAAGCTTATTTCAATGTGTGAAGTCCAA 2824
2824 ATACAGTCTCTTAAATGTGAACACCTGACATCCGAGGTTTCCACCGGTTCAAGTGGGA 2883
2825 ATACAGTCTCTTAAATGTGAACACCTGACATCCGAGGTTTCCACCGGTTCAAGTGGGA 2884
2884 GACATCCGCTTGGCATCAGCAGCAGATCCAGCTCGAGCTTTCAGCTTGGTCAACCAA 2943
2885 GACATCCGCTTGGCATCAGCAGCAGATCCAGCTCGAGCTTTCAGCTTGGTCAACCAA 2944
2944 GACGGGAGGCTTGTCTCTACGATGGAGGTCCAGCTCGGGGATCGACCTGGAGTGC 3003
2945 GACGGGAGGCTTGTCTCTACGATGGAGGTCCAGCTCGGGGATCGACCTGGAGTGC 3004
3004 ACATGCGCCCTGATGAGCAGCTTCCGCTGAGCTTGGAGGTTCAAGCATGCGCAGCTGAG 3063
3005 ACATGCGCCCTGATGAGCAGCTTCCGCTGAGCTTGGAGGTTCAAGCATGCGCAGCTGAG 3064
3064 AACAGGCTTAACTCCCTTCCACCGGCTTCCACCTGCGGCTTCACTGCGGAAAGCAGCTTCTG 3123
3065 AACAGGCTTAACTCCCTTCCACCGGCTTCCACCTGCGGCTTCACTGCGGAAAGCAGCTTCTG 3124
3124 CTGGTGCACGATGCTGCCCTGAAAAACAGGCTCAGCCGTTCCAGGGGATTCGACGCC 3183
3125 CTGGTGCACGATGCTGCCCTGAAAAACAGGCTCAGCCGTTCCAGGGGATTCGACGCC 3184
3184 CCGGCTCAGTGGGAAACAGGCTTCCGAGCAGCAAGTGGGGGCGACGAGATGCC 3243
3185 CCGGCTCAGTGGGAAACAGGCTTCCGAGCAGCAAGTGGGGGCGACGAGATGCC 3244
3244 TCCAGGATTTCACTCAGCTGAGCCCTGCGCCACCTGCTGAAAAAACATCCCGCAGTGAA 3303
3245 TCCAGGATTTCACTCAGCTGAGCCCTGCGCCACCTGCTGAAAAAACATCCCGCAGTGAA 3304
3304 GAGACAGAGGAGGATGGCAGGATTAACCTGGGAAACAAACAGGAGTCTTTTCTGCC 3363
3305 GAGACAGAGGAGGATGGCAGGATTAACCTGGGAAACAAACAGGAGTCTTTTCTGCC 3364
3364 CTTGCTCAGTGGCTTGGCTTGGATCAGTGAACCTTTGTTGGCAGACAGG 3423
3365 CTTGCTCAGTGGCTTGGCTTGGATCAGTGAACCTTTGTTGGCAGACAGG 3424
3424 GGACAGAGCTTCCAGGCTTGGGCGAGGCTTGGGCGAGGCTTCCGCTTCCAGCTC 3483
3425 GGACAGAGCTTCCAGGCTTGGGCGAGGCTTGGGCGAGGCTTCCGCTTCCAGCTC 3484
3484 CAGGCTGCTGTGAGAGTGTCAAGTGTGTGAAGGGGCCCAACTCAGGTTTCAAGTGTGAGAACCA 3543

Db	3485		CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCCAAACTTCAGTTTCAGTGCAGAAACCA	3544
Qy	3544		GGTCAGCAGAGTATGCCCGCCCGTAGTGTAAAGGGGCCCTCTAAACCCCTTGCCTGGCCCTC	3603
Db	3545		GGTCAGCAGAGTATGCCCGCCCGTAGTGTAAAGGGGCCCTCTAAACCCCTTGCCTGGCCCTC	3604
Qy	3604		ACCTGGCCAGCTCAACCCCTTTTGGGTGTAGGGGAAAAAGATGCTGACCTCGGGAAAGGCT	3663
Db	3605		ACCTGGCCAGCTCAACCCCTTTTGGGTGTAGGGGAAAAAGATGCTGACCTCGGGAAAGGCT	3664
Qy	3664		CCCTGGTAGAATACACACACTTTTCAGGTGTGTGCAACACAGGTCCTGAGTTGACCTCT	3723
Db	3665		CCCTGGTAGAATACACACACTTTTCAGGTGTGTGCAACACAGGTCCTGAGTTGACCTCT	3724
Qy	3724		GGTTTCAGCAAGAGCAAAAGAGGTGTGAAGTCAAGTGGTTCTCAGTCCCCAGACATGT	3783
Db	3725		GGTTTCAGCAAGAGCAAAAGAGGTGTGAAGTCAAGTGGTTCTCAGTCCCCAGACATGT	3784
Qy	3784		GCCCTTTTGTCTGGCTTACCACTTTTCCCGAGACAGAGGCCCGCCAGAGCCCTTTCAGGC	3843
Db	3785		GCCCTTTTGTCTGGCTTACCACTTTTCCCGAGACAGAGGCCCGCCAGAGCCCTTTCAGGC	3844
Qy	3844		CCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTGAT	3903
Db	3845		CCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTGAT	3904
Qy	3904		GCAGATATGCTTCACAGGAACAGTCTGTGATGGACATGATCAGTGCCTTAAGGAAGCAG	3963
Db	3905		GCAGATATGCTTCACAGGAACAGTCTGTGATGGACATGATCAGTGCCTTAAGGAAGCAG	3964
Qy	3964		CAGAGAGAGCTGCGGGCGCCAGCCCCACTATCAGTGTCCAGCGTGTCTGGTTCCCCAG	4023
Db	3965		CAGAGAGAGCTGCGGGCGCCAGCCCCACTATCAGTGTCCAGCGTGTCTGGTTCCCCAG	4024
Qy	4024		AGCAAGCTCAGCATCACA CTGACACTCAACCTGCTGCCCTGCCCTTGGCCAGAGGGTACTGC	4083
Db	4025		AGCAAGCTCAGCATCACA CTGACACTCAACCTGCTGCCCTGCCCTTGGCCAGAGGGTACTGC	4084
Qy	4084		CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTCGCAGGGC	4143
Db	4085		CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTCGCAGGGC	4144
Qy	4144		AGGCGAGGGCAGTGACACTGTAGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGACACA	4203
Db	4145		AGGCGAGGGCAGTGACACTGTAGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGACACA	4204
Qy	4204		GTCCTTGAGCTGTCCACATGATGTGACTCTTCAAACTCTTCCAGATTTCTCTAAGAATA	4263
Db	4205		GTCCTTGAGCTGTCCACATGATGTGACTCTTCAAACTCTTCCAGATTTCTCTAAGAATA	4264
Qy	4264		GCACCCCTTCCCATTTGCCCCAGCTTAGCCTCTTCTCCAGGGGAGCTACTCAGGACTC	4323
Db	4265		GCACCCCTTCCCATTTGCCCCAGCTTAGCCTCTTCTCCAGGGGAGCTACTCAGGACTC	4324
Qy	4324		ACGTAGCATTAATACAGCTGTGAATCTGCAGGGGTGTCTGTAGCCCTCAACCTCTCTGGG	4383
Db	4325		ACGTAGCATTAATACAGCTGTGAATCTGTAGGGGTGTCTGTAGCCCTCAACCTCTCTGGG	4384
Qy	4384		GCAGGGAGCGCCGAGACTCCGTGGGGAAGCTCATTTCCCATCTTTTCCCAAGACAGCCTT	4443
Db	4385		GCAGGGAGCGCCGAGACTCCGTGGGGAAGCTCATTTCCCATCTTTTCCCAAGACAGCCTT	4444
Qy	4444		TGTCAGGCTGTCCACATTTGAGTGTAGACTGTCTCCGGGGAGAGACCCCGGCCGCCAGCAC	4503
Db	4445		TGTCAGGCTGTCCACATTTGAGTGTAGACTGTCTCCGGGGAGAGACCCCGGCCGCCAGCAC	4504
Qy	4504		ATAAAGAACTGCAGCCTTTGGTACTGCAGAGTCTGGGTGTGTAGAGAACTCTTTGTAGCAA	4563
Db	4505		ATAAAGAACTGCAGCCTTTGGTACTGCAGAGTCTGGGTGTGTAGAGAACTCTTTGTAGCAA	4564
Qy	4564		TAAAGTTTGGGGTGATGACAAATGTTAAAAA	4595

Db 4565 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4596

RESULT 13

US-09-949-003C-1291

; Sequence 1291, Application US/09949003C

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH NEUROPSYCH

; TITLE OF INVENTION: DISORDERS, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL000791

; CURRENT APPLICATION NUMBER: US/09/949,003C

; CURRENT FILING DATE: 2001-09-10

; PRIOR APPLICATION NUMBER: 60/231,446

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 74065

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1291

; LENGTH: 4584

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-949-003C-1291

Query Match 96.6%; Score 4440; DB 41; Length 4584;

Best Local Similarity 99.1%; Pred. No. 0;

Matches 4549; Conservative 0; Mismatches 30; Indels 11; Gaps 8

Qy	3	CGGGGGGACCTGTGCCGTGTGGAA	CGTGTAGCTGTGTGA	-AGGTGGACTCTGTTACCATTGA	61
Db	3	CGGGGGGACCTGTGCCGTGTGGAA	CGTGTAGCTGTGTGA	CGTGTGGAGTCTCTGTTACCATTGA	62
Qy	62	GGATGTTTGGAGGATGATGATGTGTGG	CAGAGGCACACATAAACAGG	CAGACACCTTTTG	121
Db	63	GGATGTTTGGAGGATGATGATGTGTGG	CAGAGGCACACATAAACAGG	CAGACACCTTTTG	122
Qy	122	CCCCCTGCTTTCTCCCCCAACCCCAAG	GGTGACCTGTGTCTCC	CAGGTCGCGGATTCCTAA	181
Db	123	CCCCCTGCTTTCTCCCCCAACCCCAAG	GGTGACCTGTGTCTCC	CAGGTCGCGGATTCCTAA	182
Qy	182	GTGACCTGCTCTGTGTTTGGTCTCTCT	CAGATGAGCACACAGCTGGGAGATGG	CAGATGA	241
Db	183	GTGACCTGCTCTGTGTTTGGTCTCTCT	CAGATGAGCACACAGCTGGGAGATGG	CAGATGA	242
Qy	242	TGGAATATGGCTGCCCAGGTGCCCTTG	CTCAGCAGTGGGGCAGCAGAAGGAAC	CTCCCCA	301
Db	243	TGGAATATGGCTGCCCAGGTGCCCTTG	CTCAGCAGTGGGGCAGCAGAAGGAAC	CTCCCCA	302
Qy	302	AGCCAAAGGAGAAAGACGCGCCCACT	GGGGGAAAGAAACAGAGCTCCGTCT	CACAGCTTGAGG	361
Db	303	AAGCCAAAGGAGAAAGACGCGCCCACT	GGGGGAAAGAAACAGAGCTCCGTCT	CACAGCTTGAGG	362
Qy	362	CCGTGGAGAGAGCCCTGTGTTCTGCG	GAAGTGGGAGATCCTGAATCAGCTGAT	CAACA	421
Db	363	CCGTGGAGAGAGCCCTGTGTTCTGCG	GAAGTGGGAGATCCTGAATCAGCTGAT	CAACA	422
Qy	422	AGGSCACAGCCAAAGGAAGGCTCCG	AGGCAGGCAGCTGCCATCTCTATCAT	CGCCCCAGG	481
Db	423	AGGSCACAGCCAAAGGAAGGCTCCG	AGGCAGGCAGCTGCCATCTCTATCAT	CGCCCCAGG	482
Qy	482	CTGAGTGTGAGAAATAGCCAAAGATT	CAGCCCCACCTTTTTCAGAA	CGCATATTTTCATCGCTG	541
Db	483	CTGAGTGTGAGAAATAGCCAAAGATT	CAGCCCCACCTTTTTCAGAA	CGCATATTTTCATCGCTG	542
Qy	542	GGTCCAAACAGTACAGCCAGTCCG	AGAGTCTTGATCAGATCCCAACAA	TAATGTGGGCCATG	601
Db	543	GGTCCAAACAGTACAGCCAGTCCG	AGAGTCTTGATCAGATCCCAACAA	TAATGTGGGCCATG	602
Qy	602	CTACAGAGGGCAAAATGGCCCGTGT	GTGTTGGAAGGGGAAAGCGTCG	CAGCAAGCCCGA	661
Db	603	CTACAGAGGGCAAAATGGCCCGTGT	GTGTTGGAAGGGGAAAGCGTCG	CAGCAAGCCCGA	662
Qy	662	AGAAACGGGAAGGAAGAGCTCAA	AAGTCCCTTGGCTCATG	CAGAGAGTGGCCTTTGGCCAAAC	721

Db 663 AGAAACGGAAGAGAGCTCAAAAGTCCCTGGCTCATGCAAGGAGTGGCCCTTGCCAAAC 722
Qy 722 CCCTCCAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGCAGGAGATGAGTCTC 781
Db 723 CCCTCCAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGCAGGAGATGAGTCTC 782
Qy 782 CACTGGGGCCCCATATGTTTAAAGACACCCCGCAGTTTCAACAGGCTCTGAAGAACGAC 841
Db 783 CACTGGGGCCCCATATGTTTAAAGACACCCCGCAGTTTCAACAGGCTCTGAAGAACGAC 842
Qy 842 GCCTTGGGCAACTCTGTTTAAAGCAGCTTGGCAGGGCTTACGCGCCGCTCTGCTCGAT 901
Db 843 GCCTTGGGCAACTCTGTTTAAAGCAGCTTGGCAGGGCTTACGCGCCGCTCTGCTCGAT 902
Qy 902 CAGAACTCCAAACATGATCAGCCCTTGCATATGTGAAACAGGTGTGGAACCTGCACC 961
Db 903 CAGAACTCCAAACATGATCAGCCCTTGCATATGTGAAACAGGTGTGGAACCTGCACC 962
Qy 962 ACCGCCAGGACGGAGGCCCTTGCCCTTGCCCAAGCAGCCCTTCCCTATAGCAGACTGC 1021
Db 963 ACCGCCAGGACGGAGGCCCTTGCCCTTGCCCAAGCAGCCCTTCCCTATAGCAGACTGC 1022
Qy 1022 CTATTCCTTCCCATTCACACCTCTTCAGCCCTTGGAAACCTCACCTCTGAGTCCCTTCC 1081
Db 1023 CTATTCCTTCCCATTCACACCTCTTCAGCCCTTGGAAACCTCACCTCTGAGTCCCTTCC 1082
Qy 1082 TGGGCAAACTGGCCCTGTGTAGACAGCCAGAAACCTTGCCTGACCCACCTGAGCAAAC 1141
Db 1083 TGGGCAAACTGGCCCTGTGTAGACAGCCAGAAACCTTGCCTGACCCACCTGAGCAAAC 1142
Qy 1142 TGGCCCTGTGTAGACAGTCCAAAGCCCTTGCTGGCCCAACCTTGAGCCCAAGCTGCCTGT 1201
Db 1143 TGGCCCTGTGTAGACAGTCCAAAGCCCTTGCTGGCCCAACCTTGAGCCCAAGCTGCCTGT 1202
Qy 1202 CTCTGTGTCCATGAGAGTTTCTGTGGAGGATACCTAGTGCATGCTCTGCAAGGCA 1261
Db 1203 CTCTGTGTCCATGAGAGTTTCTGTGGAGGATACCTAGTGCATGCTCTGCAAGGCA 1262
Qy 1262 GCCTGAGCTCAAGCCAGGCCACAGCCTTGACAGCCTGGCCCAAGCCTGGGCAAGGG 1321
Db 1263 GCCTGAGCTCAAGCCAGGCCACAGCCTTGACAGCCTGGCCCAAGCCTGGGCAAGGG 1322
Qy 1322 GCTCCAGATCCGGGAGCCAGCCCAAAAATGAGGACAAAGAGGCTGCTGCTCACTG 1381
Db 1323 GCTCCAGATCCGGGAGCCAGCCCAAAAATGAGGACAAAGAGGCTGCTGCTCACTG 1382
Qy 1382 AGAACTCAAGCCAGTGTATGAGTACAGAGAGTCCACTGGGCCAGCAGCAGC 1441
Db 1383 AGAACTCAAGCCAGTGTATGAGTACAGAGAGTCCACTGGGCCAGCAGCAGC 1442
Qy 1442 TCCGCTGGGAGAGGCTCCCTTGGAGAGGTGCAAGAGTGGAGGACAAAGCAGACTGGCT 1501
Db 1443 TCCGCTGGGAGAGGCTCCCTTGGAGAGGTGCAAGAGTGGAGGACAAAGCAGACTGGCT 1502
Qy 1502 TCCAGTGGCTGTCAAAAGGTGGCCCTGGAAGTATTTTCGGGCGAGAGAGCTGATGGCAT 1561
Db 1503 TCCAGTGGCTGTCAAAAGGTGGCCCTGGAAGTATTTTCGGGCGAGAGAGCTGATGGCAT 1562
Qy 1562 GTGAGGATGACCTACCCAGAAATTTGTCCTTTGTATGAGAGTGTGAGAGAGGGCTT 1621
Db 1563 GTGAGGATGACCTACCCAGAAATTTGTCCTTTGTATGAGAGTGTGAGAGAGGGCTT 1622
Qy 1622 GGGTCAACATCTTCATGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTGTCAAGGAGC 1681
Db 1623 GGGTCAACATCTTCATGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTGTCAAGGAGC 1682
Qy 1682 AGGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTTGGGCCAGGCCCTGGAGGGTCTGG 1741
Db 1683 AGGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTTGGGCCAGGCCCTGGAGGGTCTGG 1742
Qy 1742 AATACCTCACTCAAGAGGATTTCTGCATGGGGAGCTCAAGAGCTGCAACCTGCTCTGT 1801
Db 1743 AATACCTCACTCAAGAGGATTTCTGCATGGGGAGCTCAAGAGCTGCAACCTGCTCTGT 1802

Qy 1802 CCAGCGATGGAGGCCAGCAGCCCTCTGTGACTTTTGGCCATGCTGTGTGTTTCAACCTG 1861
Db 1803 CCAGCGATGGAGGCCAGCAGCCCTCTGTGACTTTTGGCCATGCTGTGTGTTTCAACCTG 1862
Qy 1862 ATGSCCTTGGGAAAGTCTCTCACAGGGAGCTACATCCCTTGGCCACAGAGACCCACATGG 1921
Db 1863 ATGSCCTTGGGAAAGTCTCTCACAGGGAGCTACATCCCTTGGCCACAGAGACCCACATGG 1922
Qy 1922 CTCCGAGGTGGTCTGGGAGAGGCTGCGACGCAAGGTGGATGTCTGAGAGAGCTGCT 1981
Db 1923 CTCCGAGGTGGTCTGGGAGAGGCTGCGACGCAAGGTGGATGTCTGAGAGAGCTGCT 1982
Qy 1982 GTATGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCGAGGGGCGC 2041
Db 1983 GTATGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCGAGGGGCGC 2042
Qy 2042 TCTGCTCTCAAGTTCAGCAGGAGCTCCGCTGTGAGGGAGATCCCACTCTTCGCGCC 2101
Db 2043 TCTGCTCTCAAGTTCAGCAGGAGCTCCGCTGTGAGGGAGATCCCACTCTTCGCGCC 2102
Qy 2102 CTCTCAGAGCCAGGCTCATCCAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTG 2161
Db 2103 CTCTCAGAGCCAGGCTCATCCAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTG 2162
Qy 2162 CAGCGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCC 2221
Db 2163 CAGCGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCC 2222
Qy 2222 CTTGGAGGGAGATATAAAGAACCAAGACATCCACCGCCCAATCAAGCCCAATTACCAC 2281
Db 2223 CTTGGAGGGAGATATAAAGAACCAAGACATCCACCGCCCAATCAAGCCCAATTACCAC 2282
Qy 2282 AGACCTTCATGCCCAGCAGAGAGCTTTGCGCAAGGGCCCGAGGCCCCGGCCAGCTG 2341
Db 2283 AGACCTTCATGCCCAGCAGAGAGCTTTGCGCAAGGGCCCGAGGCCCCGGCCAGCTG 2342
Qy 2342 AGGAGACAACAGGCGAGAGCCCTTAAGCTCAGGCTCCTCTCCCAAGAGAGCCCGAGAGC 2401
Db 2343 AGGAGACAACAGGCGAGAGCCCTTAAGCTCAGGCTCCTCTCCCAAGAGAGCCCGAGAGC 2402
Qy 2402 CAACCAAGTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGATGTGGGAAACCTTAC 2461
Db 2403 CAACCAAGTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGATGTGGGAAACCTTAC 2462
Qy 2462 CTCTGCTCTCCCTGGAGCAGCCCTGCGCAAGAACCCAGCTCACAGAGCGGAAAGCAA 2521
Db 2463 CTCTGCTCTCCCTGGAGCAGCCCTGCGCAAGAACCCAGCTCACAGAGCGGAAAGCAA 2522
Qy 2522 CCGTCCCGAGCAGGAACTGCAGCAGCTGGAAATAGAAATATTTCTCAACAGCCTGTCCC 2581
Db 2523 CCGTCCCGAGCAGGAACTGCAGCAGCTGGAAATAGAAATATTTCTCAACAGCCTGTCCC 2582
Qy 2582 AGCCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCAGCAGCTCT 2641
Db 2583 AGCCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCAGCAGCTCT 2642
Qy 2642 CCCTGTCGAGTACAGTGAAGAACCCATCAAGGCTCTCAAGCTTCGCGGAGCAGCC 2701
Db 2643 CCCTGTCGAGTACAGTGAAGAACCCATCAAGGCTCTCAAGCTTCGCGGAGCAGCC 2702
Qy 2702 TGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCTCGAAGCTTCCAGCTTGAACA 2761
Db 2703 TGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCTCGAAGCTTCCAGCTTGAACA 2762
Qy 2762 TGAGTCTGGCCCGGGGGCGGCCCAACCAAGCTTCAATGAGTGTGAAGAGTCC 2821
Db 2763 TGAGTCTGGCCCGGGGGCGGCCCAACCAAGCTTCAATGAGTGTGAAGAGTCC 2822
Qy 2822 AATACAGTCTCTTAATGAGTGAACCTTCACATCCCGGAGTTCACCGGGTCAAGTGG 2881
Db 2823 AATACAGTCTCTTAATGAGTGAACCTTCACATCCCGGAGTTCACCGGGTCAAGTGG 2882

2882 GAGACATCCGCATGCGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGCTTGGTCAACA 2941
2883 GAGACATCCGCATGCGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGCTTGGTCAACA 2942
2942 AAGACGGGAGCCTGTTCCCTAGCAGATGGAGGTGCAGACTCGGGCATCGACCTGCACT 3001
2943 AAGACGGGAGCCTGTTCCCTAGCAGATGGAGGTGCAGACTCGGGCATCGACCTGCACT 3002
3002 GCACACTGGCCCTGATGCGACTTCGCTGAGCTGGAGGTGAAGCATGCGCAGCTGG 3061
3003 GCACACTGGCCCTGATGCGACTTCGCTGAGCTGGAGGTGAAGCATGCGCAGCTGG 3062
3062 AGAACAGGCGCTAACCCCTCCCTCCACCGCCGCTCCACACTGCGGAAAGCAGCCTTCC 3121
3063 AGAACAGGCGCTAACCCCTCCCTCCACCGCCGCTCCACACTGCGG -AAGCAGCCTTCC 3121
3122 TGCTGGGTGACGATGCTGCTGCTGAAACACAGGCTCAGCGGTTCCAGGGGATGCGCAG 3181
3122 TGCTGGGTGACGATGCTGCTGCTG -AAACACAGGCTCAGCGGTTCCAGGGGATGCGCAG 3180
3182 CCCCCGGCTCAGGTGGGAACAGGGCTTCGACAGCAAGGTGGGGCAAGCAGAAATG 3241
3181 -CCCCGGCTCAGGTGGGAACAGGGCTTCGACAGCAAGGTGGGGCAAGCAGAAATG 3239
3242 CTTCCAGGATTCACACTGAGCCTGCCCCACCTGCTGAAAAACA -TCCGCCACGT 3300
3240 CTTCCAGGATTCACACTGAGCCTGCCCCACCTGCTGAAAAACA -CACTCCGCCACGT 3299
3301 GAAGACAGAGGAGGATGGCAGAGATTACCTGGGGAAACAAACAGGGATCTTTTCT 3360
3300 GAAGACAGAG -AGGAGGATGGCAGGATTACCTCGGGAAACAAACAGG -ATCTTCTC 3354
3361 GCCCTGCTCCAGTTCAGTTCGCTGACCCGCTTGGATCAGTGACCATTTGTTGGCAGAC 3420
3355 TGCCCTGCTCCAGTTCAGTTCGCTGACCCGCTTGGATCAGTGACCATTTGTTGGCAGAC 3414
3421 AGGGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGCGAGCCTTCCGCCCTCACC 3480
3415 AGGGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGCGAGCCTTCCGCCCTCACC 3474
3481 CTCAGGCTGCTGAGAGTGAAGTGTGAAGGGCCCAACTCAGGTTTCAGTGAGAA 3540
3475 CTCAGGCTGCTGAGAGTGAAGTGTGAAGGGCCCAACTCAGGTTTCAGTGAGAA 3534
3541 CCAGGTTCAGCAGTATGCCCGCCCTGAGTTAAGGGGCCCTCTAAACCCCTTGCCTGCG 3600
3535 CCAGGTTCAGCAGTATGCCCGCCCTGAGTTAAGGGGCCCTCTAAACCCCTTGCCTGCG 3594
3601 CTCACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAGAAATGCCTGACCCCTGGGAG 3660
3595 CTCACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAGAAATGCCTGACCCCTGGGAG 3654
3661 GCTCCCTGTTAGAAATACACCACTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACC 3720
3655 GCTCCCTGTTAGAAATACACCACTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACC 3714
3721 TCTGTTTCAGCAAGACCAAGAGGTGTGTAAGTGAAGTGTCTCAGTCCCCAGACA 3780
3715 TCTGTTTCAGCAAGACCAAGAGGTGTGTAAGTGAAGTGTCTCAGTCCCCAGACA 3774
3781 TGTGCCCTTTGCTGCTGCTACCACTCTTCCCCAGAGCAGAGCCCGGAGCCCTTCA 3840
3775 TGTGCCCTTTGCTGCTGCTACCACTCTTCCCCAGAGCAGAGCCCGGAGCCCTTCA 3834
3841 GGGCCAGACCTGCCCCAGACTGCTGGCACTCAGTTCCCTGATCTGTAAGGTGAAGGT 3900
3835 GGGCCAGACCTGCCCCAGACTGCTGGCACTCAGTTCCCTGATCTGTAAGGTGAAGGT 3894
3901 GATGAGGATATGCTCTGACAGAAACAGTCTGTGGATGGACATGATCAGTCTTAAGGAAAG 3960
3895 GATGAGGATATGCTCTGACAGAAACAGTCTGTGGATGGACATGATCAGTCTTAAGGAAAG 3954
3961 CAGCAGAGAGAGAGC -TCCGGCGCCCCAGCCCCCACTATCAGTGTCCAGCGTGTGCTGTTCC 4019

3955 CAGCAGAGAGAGAGCCTCCGGCGCCCCAGCCCCCACTATCAGTGTCCAGCGTGTGTTCC 4014
4020 CCAGAGCAGCAGCTCAGCATCAGCTGACACTGACACTCACCTGCCCCCTGCCCCAGAGGTA 4079
4015 CCAGAGCAGCAGCTCAGCATCAGCTGACACTCACCTGCCCCCTGCCCCAGAGGTA 4074
4080 CTGCCAGCGGCACTTTGCACTCTGATGACCTCAAAAGCACTTTTCATGGCTGCCCTCTGGCA 4139
4075 CTGCCAGCGGCACTTTGCACTCTGATGACCTCAAAAGCACTTTTCATGGCTGCCCTCTGGCA 4134
4140 GGGCAGGCGAGGCGAGTGCACCTGAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGA 4199
4135 GGGCAGGCGAGGCGAGTGCACCTGAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGA 4194
4200 CACAGTCTTGAGCTGCTCAGCATGATGATCTCCTCAAACTCTTCCAGATTTCTCTAAG 4259
4195 CACAGTCTTGAGCTGCTCAGCATGATGATCTCCTCAAACTCTTCCAGATTTCTCTAAG 4254
4260 AATAGCACCCCTTCCCATTTGCCAGCTTAGCCTTCTCCAGGGGAGCTACTCAGG 4319
4255 AATAGCACCCCTTCCCATTTGCCAGCTTAGCCTTCTCTCCAGGGGAGCTACTCAGG 4314
4320 ACTCAGTAGCATTAATCAGCTGTGAATCGTCAAGGGGTGTCTGCTAGCCTCAACCTCC 4379
4315 ACTCAGTAGCATTAATCAGCTGTGAATCGTCAAGGGGTGTCTGCTAGCCTCAACCTCC 4374
4380 TGGGCGAGGCGAGCGGAGACTCCGTGGGAGAGTCAATCCACATCTTGCACAGACAG 4439
4375 TGGGCGAGGCGAGCGGAGACTCCGTGGGAGAGTCAATCCACATCTTGCACAGACAG 4434
4440 CTTTGTCCAGCTGTCACATTTGAGTCAAGCTGCTCCGGGAGAGAGCCCGGCCCA 4499
4435 CTTTGTCCAGCTGTCACATTTGAGTCAAGCTGCTCCGGGAGAGAGCCCGGCCCA 4494
4500 GCACATAAAGAACTGCAGCCTTGGTGTGTCAGAGTCTGGGTGTAGAGAACTCTTTGTAA 4559
4495 GCACATAAAGAACTGCAGCCTTGGTGTGTCAGAGTCTGGGTGTAGAGAACTCTTTGTAA 4554
4560 GCAATAAAGTTTGGGTGATGACAAATGTT 4589
4555 GCAATAAAGTTTGGGTGATGACAAATGTT 4584

RESULT 14
US-10-087-192-887
; Sequence 887, Application US/10087192
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 887
; LENGTH: 4684
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-887

Query Match 96.6%; Score 4440; DB 47; Length 4684;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 4549; Conservative 0; Mismatches 30; Indels 11; Gaps 8;
QY 3 CCGGGGAGCTGCGGTGGAACGTGTGTTGA-AGTGGACTGTGTACATTGA 61

Db 53 CCGGGGACTGTGCGGTGGAA CGTGTAGCTGTTGAGAGGTGACTCTGTTCACATTGA 112
Qy 62 GGATGTTTGGAGGATGAGTATGTGTGGCAGAGGCACACATAAAGCAGGAGACCTTTTG 121
Db 113 GGATGTTTGGAGGATGAGTATGTGTGGCAGAGGCACACATAAAGCAGGAGACCTTTTG 172
Qy 122 CCCCTGCTTCTCCCAACCCCAAGGCTGACCTGTGTCTCCAGGTCTGGGATTTCTAA 181
Db 173 CCCCTGCTTCTCCCAACCCCAAGGCTGACCTGTGTCTCCAGGTCTGGGATTTCTAA 232
Qy 182 GTGACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAGACCTGGGAGATGGCAGTGA 241
Db 233 GTGACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAGACCTGGGAGATGGCAGTGA 292
Qy 242 TGGAAATGGCTGCTGCCAGGTGCCCTGCTCAGCAGTGGGGCAGCAGAAGAACTCCCA 301
Db 293 TGGAAATGGCTGCTGCCAGGTGCCCTGCTCAGCAGTGGGGCAGCAGAAGAACTCCCA 352
Qy 302 AGCCAAAGGAGAAAGCCGCTGCTGGGGAAGAAACAGAGTCCGTCTACAGCTTGAGG 361
Db 353 AAGCCAAAGGAGAAAGCCGCTGCTGGGGAAGAAACAGAGTCCGTCTACAGCTTGAGG 412
Qy 362 CCGTGGAGAGAGCCCTGTGTCTCGGAAAGTGGAGATCCTGAATGACGTGATTACCA 421
Db 413 CCGTGGAGAGAGCCCTGTGTCTCGGAAAGTGGAGATCCTGAATGACGTGATTACCA 472
Qy 422 AGGACACAGCCAAAGGAGCTCCGAGGAGGCGCAGCTGCCATCTCTATCATGCCCAGG 481
Db 473 AGGACACAGCCAAAGGAGCTCCGAGGAGGCGCAGCTGCCATCTCTATCATGCCCAGG 532
Qy 482 CTGAGTGTGAGAAATAGCAAGATTCAGCCCACTTTTCAGAACGATTTTCATCGGTG 541
Db 533 CTGAGTGTGAGAAATAGCAAGATTCAGCCCACTTTTCAGAACGATTTTCATCGGTG 592
Qy 542 GGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAATGTGCCCATG 601
Db 593 GGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAATGTGCCCATG 652
Qy 602 CTACAGAGGGCAAAATGGCCCGTGTGTGTTGGAGGGAAAGCGTCGAGCAAAAGCCCGGA 661
Db 653 CTACAGAGGGCAAAATGGCCCGTGTGTGTTGGAGGGAAAGCGTCGAGCAAAAGCCCGGA 712
Qy 662 AGAAACGGAAGAAGAGCTCAAAGTCCCTGGCTCATGACAGAGTGGCTTGGCCAAAC 721
Db 713 AGAAACGGAAGAAGAGCTCAAAGTCCCTGGCTCATGACAGAGTGGCTTGGCCAAAC 772
Qy 722 CCCTCCCAAGACCCCTGAGCAGAGAGCTGCACCATCCAGTCCAGAGAGATGATCTC 781
Db 773 CCCTCCCAAGACCCCTGAGCAGAGAGCTGCACCATCCAGTCCAGAGAGATGATCTC 832
Qy 782 CACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAAGCTCTGAAGAAACAG 841
Db 833 CACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAAGCTCTGAAGAAACAG 892
Qy 842 GCCTTGGGCAACTCTGTTTAAAGCAGCTTGGCAGGCGCTACGCGCGCTCTGCTCGAT 901
Db 893 GCCTTGGGCAACTCTGTTTAAAGCAGCTTGGCAGGCGCTACGCGCGCTCTGCTCGAT 952
Qy 902 CAGAACTCCACAACTGATCAGCCCTTGCAATGTCTGAACCAAGTGTGAAACTGCAAC 961
Db 953 CAGAACTCCACAACTGATCAGCCCTTGCAATGTCTGAACCAAGTGTGAAACTGCAAC 1012
Qy 962 ACCCCAGGAGGAGGCGCCCTGCTGCCACGCAACCCCTTCCCTATAGCAGACTGC 1021
Db 1013 ACCCCAGGAGGAGGCGCCCTGCTGCCACGCAACCCCTTCCCTATAGCAGACTGC 1072
Qy 1022 CTATATCCCTTCCCAATTCACCCCTCTCCAGCCCTTGGAAACCTCACCTCTGGAGTCTTCC 1081
Db 1073 CTATATCCCTTCCCAATTCACCCCTCTCCAGCCCTTGGAAACCTCACCTCTGGAGTCTTCC 1132
Qy 1082 TGGGCAAACTGGCCTGTGTAGACAGCCAGAAACCCCTTGGCTGACCCACACTGAGAAAC 1141
Db 1133 TGGGCAAACTGGCCTGTGTAGACAGCCAGAAACCCCTTGGCTGACCCACACTGAGAAAC 1192

Qy 1142 TGGCCTGTGTAGACAGTCCAAAAGCCCTGTGCTGCCCCACACCTGGAGCCCAAGCTGCTGT 1201
Db 1193 TGGCCTGTGTAGACAGTCCAAAAGCCCTGTGCTGCCCCACACCTGGAGCCCAAGCTGCTGT 1252
Qy 1202 CTGTGTGTGCCCATGAGAGTCTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGCA 1261
Db 1253 CTGTGTGTGCCCATGAGAGTCTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGCA 1312
Qy 1262 GCCTGAGCTCAAGCAGGCCCCACAGCCTGACCAAGCTGGCCCAAGACCTGGGAGCAGCGG 1321
Db 1313 GCCTGAGCTCAAGCAGGCCCCACAGCCTGACCAAGCTGGCCCAAGACCTGGGAGCAGCGG 1372
Qy 1322 GCTCCAGATCCCGGAGGCCCCAGGCCCAAAACTGAGGACAACGAGGGTGTCTGCTCACTG 1381
Db 1373 GCTCCAGATCCCGGAGGCCCCAGGCCCAAAACTGAGGACAACGAGGGTGTCTGCTCACTG 1432
Qy 1382 AGAAACTCAAGCCAGTGGATTATGAGTACCCGAGGAAGTCCACTGGGCGCACGACGAGC 1441
Db 1433 AGAAACTCAAGCCAGTGGATTATGAGTACCCGAGGAAGTCCACTGGGCGCACGACGAGC 1492
Qy 1442 TCCGCTGGCAGAGGCTCCTTCGGAGAGGTGCAACAGGATGGAGGACAAGCAGACTGGCT 1501
Db 1493 TCCGCTGGCAGAGGCTCCTTCGGAGAGGTGCAACAGGATGGAGGACAAGCAGACTGGCT 1552
Qy 1502 TCCAGTGCCTGTCAAAGAGGTGCGCTGGAAGTATTTTCGGGCGAGAGAGCTGATGGCAT 1561
Db 1553 TCCAGTGCCTGTCAAAGAGGTGCGCTGGAAGTATTTTCGGGCGAGAGAGCTGATGGCAT 1612
Qy 1562 GTGAGAGATTGACCTCACCCAGAAATTTGTCTCTTTGTATGAGAGCTGTGAGAGAGGGCTT 1621
Db 1613 GTGAGAGATTGACCTCACCCAGAAATTTGTCTCTTTGTATGAGAGCTGTGAGAGAGGGCTT 1672
Qy 1622 GGGTCAACATCTTCATGGAGCTGCTGGAAGTGTCTCCCTGGGCGAGCTGTGCAAGGAGC 1681
Db 1673 GGGTCAACATCTTCATGGAGCTGCTGGAAGTGTCTCCCTGGGCGAGCTGTGCAAGGAGC 1732
Qy 1682 AGGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTTCCTGGGCGAGGCCCTGGAGGGTCTGG 1741
Db 1733 AGGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTTCCTGGGCGAGGCCCTGGAGGGTCTGG 1792
Qy 1742 AATACCTCCTACTCAGAGAGATTTGTCATGGGAGCTCAAGAGCTGACAACTGCTCTGT 1801
Db 1793 AATACCTCCTACTCAGAGAGATTTGTCATGGGAGCTCAAGAGCTGACAACTGCTCTGT 1852
Qy 1802 CCAGCAGTGGAGCCACGAGCCCTGTGACTCTTGGCCATGCTGTGCTCTTCAACCTG 1861
Db 1853 CCAGCAGTGGAGCCACGAGCCCTGTGACTCTTGGCCATGCTGTGCTCTTCAACCTG 1912
Qy 1862 ATGGCTTGGGAAAGTCTTGTCTCACAGGGACTACATCCCTGGCAGAGACCCACATGG 1921
Db 1913 ATGGCTTGGGAAAGTCTTGTCTCACAGGGACTACATCCCTGGCAGAGACCCACATGG 1972
Qy 1922 CTCGAGAGTGTGTCTGGGAGAGCTGCGACGCAAGGTGGATGTCTGAGAGAGCTGCT 1981
Db 1973 CTCGAGAGTGTGTCTGGGAGAGCTGCGACGCAAGGTGGATGTCTGAGAGAGCTGCT 2032
Qy 1982 GTATGATGTCACATGCTCAACGGCTGCCACCCCTGGAATCAGTTCTTCGAGGGCGC 2041
Db 2033 GTATGATGTCACATGCTCAACGGCTGCCACCCCTGGAATCAGTTCTTCGAGGGCGC 2092
Qy 2042 TCTGCTCTCAAGATTTGCGAGGAGCTCCGCTGTGAGGAGATCCCACTCTCTGCGGCC 2101
Db 2093 TCTGCTCTCAAGATTTGCGAGGAGCTCCGCTGTGAGGAGATCCCACTCTCTGCGGCC 2152
Qy 2102 CTCTCAGCCCGAGGCGCATCAAAGAGGGGTGAGGAAAGAGCCCATCAACGCGCTGTCTG 2161
Db 2153 CTCTCAGCCCGAGGCGCATCAAAGAGGGGTGAGGAAAGAGCCCATCAACGCGCTGTCTG 2212
Qy 2162 CAGGGAGCTGGGAGGAGGTGAAACGGGCACTACAGAGAGTGGAGGCTCTGAGAGCC 2221
Db 2213 CAGGGAGCTGGGAGGAGGTGAAACGGGCACTACAGAGAGTGGAGGCTCTGAGAGCC 2272

QY 2222 CTTGGAGGGGAGATATATAAGAACCAAGACATCCACCGCAAAATCAAGCAATTAACACC 2281
DB 2273 CTTGGAGGGGAGATATATAAGAACCAAGACATCCACCGCAAAATCAAGCAATTAACACC 2332
QY 2282 AGACCTTCATGCCAGCCGAGAGAGCTTTTCGCAAGGCGCCCAAGGCGCCGCGGAGCTG 2341
DB 2333 AGACCTTCATGCCAGCCGAGAGAGCTTTTCGCAAGGCGCCCAAGGCGCCGCGGAGCTG 2392
QY 2342 AGGAGACACAGCGAGAGCCCTTACGCTCAGGCTCCTCTCCACCAAGAGCCGCCAGAGC 2401
DB 2393 AGGAGACACAGCGAGAGCCCTTACGCTCAGGCTCCTCTCCACCAAGAGCCGCCAGAGC 2452
QY 2402 CAAACAAGTCTCTCCCTTGAATTTGAGCAAGGAGAGTCTGGGATGTGGGAACCTTTAC 2461
DB 2453 CAAACAAGTCTCTCCCTTGAATTTGAGCAAGGAGAGTCTGGGATGTGGGAACCTTTAC 2512
QY 2462 CTCTGTCTCTCTGGAGCAGGCGCCCTTGCAGAAACCCAGCTCAACAGAGCGGAAGCAA 2521
DB 2513 CTCTGTCTCTCTGGAGCAGGCGCCCTTGCAGAAACCCAGCTCAACAGAGCGGAAGCAA 2572
QY 2522 CGGTCCCGAGCAGGAATCGACGAGCTGGAATAGATTTATCTCAACAGCCTGTCC 2581
DB 2573 CGGTCCCGAGCAGGAATCGACGAGCTGGAATAGATTTATCTCAACAGCCTGTCC 2632
QY 2582 AGCCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCCTCT 2641
DB 2633 AGCCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCCTCT 2692
QY 2642 CCCTGTCTGATGACAGTAGAGAGAAACCCATCAAGGCTCTCAAGCTCGGGGACACC 2701
DB 2693 CCCTGTCTGATGACAGTAGAGAGAAACCCATCAAGGCTCTCAAGCTCGGGGACACC 2752
QY 2702 TGAGCTCAGGCGTACACTCTCTGAGCAGCAGGCGGAGCTCGAAGCTCAGCTGGAACA 2761
DB 2753 TGAGCTCAGGCGTACACTCTCTGAGCAGCAGGCGGAGCTCGAAGCTCAGCTGGAACA 2812
QY 2762 TGGTGTCTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAGTCC 2821
DB 2813 TGGTGTCTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAGTCC 2872
QY 2822 AAATACAGTCTCTTAATGTGTGAACCACTTGCAATCCGGGAGTTCAACCGGGTCAAAGTGG 2881
DB 2873 AAATACAGTCTCTTAATGTGTGAACCACTTGCAATCCGGGAGTTCAACCGGGTCAAAGTGG 2932
QY 2882 GAGACATCGCCTGACATGACGACGATCCAGCTGCAGCTTCAAGCTTGTGTGACCA 2941
DB 2933 GAGACATCGCCTGACATGACGACGATCCAGCTGCAGCTTCAAGCTTGTGTGACCA 2992
QY 2942 AAGACGGGAGGCTGTCTGCTACGACATGGAGGTGCCAGCTCGGGCATTCGACCTGCAAT 3001
DB 2993 AAGACGGGAGGCTGTCTGCTACGACATGGAGGTGCCAGCTCGGGCATTCGACCTGCAAT 3052
QY 3002 GACACTGCGCCCTGATGCGAGCTTGCCTTGAGCTGGAGGTCAAGCATGCGGCGAGCTGG 3061
DB 3053 GACACTGCGCCCTGATGCGAGCTTGCCTTGAGCTGGAGGTCAAGCATGCGGCGAGCTGG 3112
QY 3062 AGAAAGGCGCTTAAACCTGCTTCAACCGCGGCTTCAACCTGCGGGAAGCAGCCTTCC 3121
DB 3113 AGAAAGGCGCTTAAACCTGCTTCAACCGCGGCTTCAACCTGCGGGAAGCAGCCTTCC 3171
QY 3122 TGCTCGGTGACGATGCTGCTGCTGAAACACAGGCTCAGCCTGTCCAGGGGATTCGCAAG 3181
DB 3172 TGCTCGGTGACGATGCTGCTGCTGAAACACAGGCTCAGCCTGTCCAGGGGATTCGCAAG 3230
QY 3182 CCCCCCGGTCAAGTGGGAACAGGCGCTTCGACGACGAGGTGGGGCAAGCAGAATG 3241
DB 3231 -CCCCCGGTCAAGTGGGAACAGGCGCTTCGACGACGAGGTGGGGCAAGCAGAATG 3289
QY 3242 CTTCCCGGATTTCACTGAGCGCTTCCCGACCTCTGCTGAAACAACTCCGCGCAGT 3300
DB 3290 CTTCCCGGATTTCACTGAGCGCTTCCCGACCTCTGCTGAAACAACTCCGCGCAGT 3349
QY 3301 GAAGAGACAGAGGAGGATGGGAGGATTTACCTGGGGAAACAAACAGGGATCTTTTCT 3360

DB 3350 GAAGAGACAG-AGGAGGATGGCAGGATTAACCTCGGGAACAAACAGG----ATCTTCTC 3404
QY 3361 GCCCTGTCTCAGTCCAGTTCGCTGACCCGCTTGGATCAGTGACATTTGTGTGGCAGAC 3420
DB 3405 TGCCCTGTCTCCAGTCCAGTTCGCTGACCCGCTTGGATCAGTGACATTTGTGTGGCAGAC 3464
QY 3421 AGGGAGAGCAGCTTCCAGCCTGGGTCAGAGGGGTGGGCGAGCCTTCGGGCCCTCACC 3480
DB 3465 AGGGAGAGCAGCTTCCAGCCTGGGTCAGAGGGGTGGGCGAGCCTTCGGGCCCTCACC 3524
QY 3481 CTCACAGGCTGTCTGAGAGTGTCAAGTGTGTAAAGGCGCCAAACTCAGGTTTCAGTCAGAA 3540
DB 3525 CTCACAGGCTGTCTGAGAGTGTCAAGTGTGTAAAGGCGCCAAACTCAGGTTTCAGTCAGAA 3584
QY 3541 CCAGTCCAGCAGGATATGCCCGCTGAGGTTAAAGGGGGCCTCTAAAACCTTGGCTGGC 3600
DB 3585 CCAGTCCAGCAGGATATGCCCGCTGAGGTTAAAGGGGGCCTCTAAAACCTTGGCTGGC 3644
QY 3601 CTCACCTGCGCAGCTCACCCCTTTTGGGTGTAGGGGAAAGAAATCGCTGACCTCGGAAG 3660
DB 3645 CTCACCTGCGCAGCTCACCCCTTTTGGGTGTAGGGGAAAGAAATCGCTGACCTCGGAAG 3704
QY 3661 GCTCCTCTGTAGAAATACACACACTTTTTCAGGTTGTGCAACACAGGTCCTGAGTTGACC 3720
DB 3705 GCTCCTCTGTAGAAATACACACACTTTTTCAGGTTGTGCAACACAGGTCCTGAGTTGACC 3764
QY 3721 TCTGTTTCAGCCCAAGGACAAAGAGGTGTGTAGTGAAGTGTCTCAGTCCCCAGACA 3780
DB 3765 TCTGTTTCAGCCCAAGGACAAAGAGGTGTGTAGTGAAGTGTCTCAGTCCCCAGACA 3824
QY 3781 TGTGCCCCCTTGTGCTGCTGCTACCACTCTTCCAGAGCAGCAGGCCCGAGCCCTTCA 3840
DB 3825 TGTGCCCCCTTGTGCTGCTGCTACCACTCTTCCAGAGCAGCAGGCCCGAGCCCTTCA 3884
QY 3841 GGCCAGCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCTCATCTGTAAAAGTGAAGGT 3900
DB 3885 GGCCAGCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCTCATCTGTAAAAGTGAAGGT 3944
QY 3901 GATCAGGATATGCTGTGACAGGACAGTCTGTGGATGGACATGATCAGTGTCTAAGAAAG 3960
DB 3945 GATCAGGATATGCTGTGACAGGACAGTCTGTGGATGGACATGATCAGTGTCTAAGAAAG 4004
QY 3961 CAGCAGAGAGAGAGC-TCCGGCGCCAGCCCACTATCAGTGTCCAGCTGCTGCTGCTCC 4019
DB 4005 CAGCAGAGAGAGAGCCTCCGGCGCCAGCCCACTATCAGTGTCCAGGCTGCTGCTGCTCC 4064
QY 4020 CCAGAGCAGCTCAGCATCACTGACACACCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 4079
DB 4065 CCAGAGCAGCTCAGCATCACTGACACACCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 4124
QY 4080 GTCCGAGCGGCACTTTTGCATCTGTGATGACCTCAAGACATTTTCATGGCTGCCCTTGGCA 4139
DB 4125 GTCCGAGCGGCACTTTTGCATCTGTGATGACCTCAAGACATTTTCATGGCTGCCCTTGGCA 4184
QY 4140 GGGCAGGCGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4199
DB 4185 GGGCAGGCGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4244
QY 4200 CACAGCTTTGAGCTGTCCACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4259
DB 4245 CACAGCTTTGAGCTGTCCACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4304
QY 4260 AATAGCACCCCTTCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4319
DB 4305 AATAGCACCCCTTCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4364
QY 4320 ACTCAGTAGGATTAATCAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4379
DB 4365 ACTCAGTAGGATTAATCAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4424
QY 4380 TGGGCGAGGAGCGCCGAGACTCCGTGGGAGAGCTCAATCCACATCTTGCACAGACAG 4439

Db 4425 TGGGCGAGGAGCGCGAGACTCCGTGGGAGAGCTCATTTCCACATCTTGGCCAGACAG 4484
Qy 4440 CTTTGTCCAGCTGTCCACATTTAGTTCAGACTGCTCCCGGGAGAGAGCCCGCCGCCCA 4499
Db 4485 CTTTGTCCAGCTGTCCACATTTAGTTCAGACTGCTCCCGGGAGAGAGCCCGCCGCCCA 4544
Qy 4500 GCACATAAGAACTGCGAGCCTTGTACTGCGAGACTCTGGGTGTAGAGAACTCTTTGTAA 4559
Db 4545 GCACATAAGAACTGCGAGCCTTGTACTGCGAGACTCTGGGTGTAGAGAACTCTTTGTAA 4604
Qy 4560 GCAATAAAGTTTGGGTGATGACAAATGTT 4589
Db 4605 GCAATAAAGTTTGGGTGATGACAAATGTT 4634

RESULT 15
US-10-170-235-9152
; Sequence 9152 Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 9152
; LENGTH: 4684
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-9152

Query Match 96.6%; Score 4440; DB 49; Length 4684;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 4549; Conservative 0; Mismatches 30; Indels 11; Gaps 8;

Qy 3 CGGGGGAGCTGTGCGGTGTGAAAGCTGTAGCTGTGA-AGGTGAGCTGTGTACCATGA 61
Db 53 CGGGGGAGCTGTGCGGTGTGAAAGCTGTAGCTGTGAAGAGGTGAGCTGTGTACCATGA 112
Qy 62 GGATGTTTGGAGATGATGTTGTGGCAGAGGACACATATAAGCAGAGAGCCCTTTG 121
Db 113 GGATGTTTGGAGATGATGTTGTGGCAGAGGACACATATAAGCAGAGAGCCCTTTG 172
Qy 122 CCCCTGCTTCTCCCAAGCTGACCTGTGTCTCCAGGTCTGGGATTTCTAA 181
Db 173 CCCCTGCTTCTCCCAAGCTGACCTGTGTCTCCAGGTCTGGGATTTCTAA 232
Qy 182 GTGACCTGCTGTGTGTGTGTCTCTCAGGATGAGCACAAGCTGGGAGATGGCAGTGA 241
Db 233 GTGACCTGCTGTGTGTGTGTCTCTCAGGATGAGCACAAGCTGGGAGATGGCAGTGA 292
Qy 242 TGGAAATGGCTGCGCAGGTGCTGCTGCTCAGAGTGGGGCAGCAGAGAACTCCCA 301
Db 293 TGGAAATGGCTGCGCAGGTGCTGCTGCTCAGAGTGGGGCAGCAGAGAACTCCCA 352
Qy 302 AGCCAAAGGAGAGAGCGCCCACTGGGGAAGAAACAGAGCTCTCAGAGCTTGGG 361
Db 353 AAGCCAAAGGAGAGAGCGCCCACTGGGGAAGAAACAGAGCTCTCAGAGCTTGGG 412
Qy 362 CCGTGGAGAGAGAGCGCCCTGTGTCTCGGAAAGTGGGAGATCCTGAATGAGCTATTACCA 421
Db 413 CCGTGGAGAGAGAGCGCCCTGTGTCTCGGAAAGTGGGAGATCCTGAATGAGCTATTACCA 472
Qy 422 AGGCGACAGCCAAAGAGGCTCCAGGAGGCGCAGCTGCCATCTCTATCATGCCCCAGG 481
Db 473 AGGCGACAGCCAAAGAGGCTCCAGGAGGCGCAGCTGCCATCTCTATCATGCCCCAGG 532
Qy 482 CTGAGTGTGAATAGCCAGAGTTACGCCCACTTTTCAGAACGATTTTCATCGCTG 541
Db 533 CTGAGTGTGAATAGCCAGAGTTACGCCCACTTTTCAGAACGATTTTCATCGCTG 592
Qy 542 GGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGTATCAGATCCCCCAACAAATGTGGCCCATG 601

Db 593 GGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGCCCATG 652
Qy 602 CTACAGAGGCAAAATGGCCCGTGTGTGTGGAGGAAAGCGTCGACGAAACCCCGGA 661
Db 653 CTACAGAGGCAAAATGGCCCGTGTGTGTGGAGGAAAGCGTCGACGAAACCCCGGA 712
Qy 662 AGAAACGGAAGAGAGAGCTCAAAAGTCCCTGTGCTCATGAGAGAGTGGCTTGGCCAAAC 721
Db 713 AGAAACGGAAGAGAGAGCTCAAAAGTCCCTGTGCTCATGAGAGAGTGGCTTGGCCAAAC 772
Qy 722 CCCTCCCGCAGGACCCCTGAGCAGAGAGCTGCACATCCAGTGCAGAGAGTGCATCTC 781
Db 773 CCCTCCCGCAGGACCCCTGAGCAGAGAGCTGCACATCCAGTGCAGAGAGTGCATCTC 832
Qy 782 CACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCCTCTCAAGGAACACAG 841
Db 833 CACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCCTCTCAAGGAACACAG 892
Qy 842 GCCTTGGGCAACTCTGTTTAAAGCAGCTTGGGAGGCGCTACGCGCGCTCTGCTCGAT 901
Db 893 GCCTTGGGCAACTCTGTTTAAAGCAGCTTGGGAGGCGCTACGCGCGCTCTGCTCGAT 952
Qy 902 CAGAACTCCAAACACTGATCAGCCCTTGAATGCTCTGAACCACTGTGGAACCTGCACC 961
Db 953 CAGAACTCCAAACACTGATCAGCCCTTGAATGCTCTGAACCACTGTGGAACCTGCACC 1012
Qy 962 ACCCCAGAGAGGAGGCGCCCTGCGCTGCGCAGCAGCACCCTTCCCTATAGCAGACTGC 1021
Db 1013 ACCCCAGAGAGGAGGCGCCCTGCGCTGCGCAGCAGCACCCTTCCCTATAGCAGACTGC 1072
Qy 1022 CTATCCCTTCCATCCACCTCTCCAGCCTTGGAAACCTCACCTCTGAGAGTCTTTC 1081
Db 1073 CTATCCCTTCCATCCACCTCTCCAGCCTTGGAAACCTCACCTCTGAGAGTCTTTC 1132
Qy 1082 TGGGCAAACTGGCTGTGTAGACAGCCAGAAACCTTGGCTGACCCACACCTGAGCAAC 1141
Db 1133 TGGGCAAACTGGCTGTGTAGACAGCCAGAAACCTTGGCTGACCCACACCTGAGCAAC 1192
Qy 1142 TGGCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGAGGCCAGCTGCTGT 1201
Db 1193 TGGCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGAGGCCAGCTGCTGT 1252
Qy 1202 CTGTGTGCTCCATGAGAGTGTGTGTGGAGGAATACCTAGTGCATGCTCTGCAAGSCA 1261
Db 1253 CTGTGTGCTCCATGAGAGTGTGTGTGGAGGAATACCTAGTGCATGCTCTGCAAGSCA 1312
Qy 1262 GCGTGTGCTCAAGCAGGCGCCACAGCTGACAGCCTGGCCCAAGACCTTGGGAGCAGCGG 1321
Db 1313 GCGTGTGCTCAAGCAGGCGCCACAGCTGACAGCCTGGCCCAAGACCTTGGGAGCAGCGG 1372
Qy 1322 GCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAAGAGGTGCTGCTGCTCACTG 1381
Db 1373 GCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAAGAGGTGCTGCTGCTCACTG 1432
Qy 1382 AGAACTCAAGCCAGTGGATTTATGAGTACCGAGAAAGTCCACTGGGCCACGACCAAGC 1441
Db 1433 AGAACTCAAGCCAGTGGATTTATGAGTACCGAGAAAGTCCACTGGGCCACGACCAAGC 1492
Qy 1442 TCCGCTTGGGAGAGGCTCTCTTGGAGAGTGTGACAGGATGGAGGACAGCAGACTGGCT 1501
Db 1493 TCCGCTTGGGAGAGGCTCTCTTGGAGAGTGTGACAGGATGGAGGACAGCAGACTGGCT 1552
Qy 1502 TCCAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTTCGGGAGAGAGGCTGATGCTAT 1561
Db 1553 TCCAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTTCGGGAGAGAGGCTGATGCTAT 1612
Qy 1562 GTGAGGATTTGACCTCACCCAGGATTTGCTCCCTTTGTATGAGAGTGTGAGAAAGGCTT 1621
Db 1613 GTGAGGATTTGACCTCACCCAGGATTTGCTCCCTTTGTATGAGAGTGTGAGAAAGGCTT 1672
Qy 1622 GGGTCAAACATCTTATGAGAGTGTGGAAGGTGCTCCCTGGGCGAGCTGCTGCTCAAGGAGC 1681

Db 1673 GGGTCAACATCTTCAATGGAGCTGCTGGAAGGTGGCTCCCTGGGCCAGAGTGGTCAAGGAGC 1732
Qy 1682 AGGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGCCAGAGCCCTGGAGGGTCTGG 1741
Db 1733 AGGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGCCAGAGCCCTGGAGGGTCTGG 1792
Qy 1742 AATACCTCCACTCAGAAAGATTTCTGCAATGGGAGCTCAAAAGCTGACAACTGCTCTCTGT 1801
Db 1793 AATACCTCCACTCAGAAAGATTTCTGCAATGGGAGCTCAAAAGCTGACAACTGCTCTCTGT 1852
Qy 1802 CCAGGATGGGAGCACGAGCCCTCTGTGACTTTGGCCATGCTGTGTGCTTCAACCTG 1861
Db 1853 CCAGGATGGGAGCACGAGCCCTCTGTGACTTTGGCCATGCTGTGTGCTTCAACCTG 1912
Qy 1862 ATGGCTGGGAAAGTCTTGTCTCAGAGGGAGTACATCCCTGGCACAGAGCCCAATGG 1921
Db 1913 ATGGCTGGGAAAGTCTTGTCTCAGAGGGAGTACATCCCTGGCACAGAGCCCAATGG 1972
Qy 1922 CTCCGAGGTGGTGGGACGAGCTGCGAGCCCAAGGTGATGTCTGGAGCAGTGTCT 1981
Db 1973 CTCCGAGGTGGTGGGACGAGCTGCGAGCCCAAGGTGATGTCTGGAGCAGTGTCT 2032
Qy 1982 GTATGATGTCATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGC 2041
Db 2033 GTATGATGTCATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGC 2092
Qy 2042 TGTGCTCAAGATTGCCAGGAGCTCCGCTGTGAGGAGATCCACCCCTCTCGGCC 2101
Db 2093 TGTGCTCAAGATTGCCAGGAGCTCCGCTGTGAGGAGATCCACCCCTCTCGGCC 2152
Qy 2102 CTCTCACAGCCAGGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTG 2161
Db 2153 CTCTCACAGCCAGGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTG 2212
Qy 2162 CAGCGAGCTGGAGGAGGAGTGAACCGGCACTACAGCAAGTGGAGTCTGAAGAGCC 2221
Db 2213 CAGCGAGCTGGAGGAGGAGTGAACCGGCACTACAGCAAGTGGAGTCTGAAGAGCC 2272
Qy 2222 CTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAATCAAGCAATTAACACC 2281
Db 2273 CTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAATCAAGCAATTAACACC 2332
Qy 2282 AGACCTCTCATGCCAGCCGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGCCAGCTG 2341
Db 2333 AGACCTCTCATGCCAGCCGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGCCAGCTG 2392
Qy 2342 AGGAGACAAAGGAGAGCCCTAAGCTCCAGCTCTCTCCACAGAGCCCGCAGAGC 2401
Db 2393 AGGAGACAAAGGAGAGCCCTAAGCTCCAGCTCTCTCCACAGAGCCCGCAGAGC 2452
Qy 2402 CAAAAGTCTCTCCCTTGAATTTGAGCAAGGAGGCTCTGGGATGTGGGAACCTTAC 2461
Db 2453 CAAAAGTCTCTCCCTTGAATTTGAGCAAGGAGGCTCTGGGATGTGGGAACCTTAC 2512
Qy 2462 CTCTGTCTCTCCCTGGAGCCAGCCCTGCCAGAAACCCAGCTCAAGAGCGGAAAGCAA 2521
Db 2513 CTCTGTCTCTCCCTGGAGCCAGCCCTGCCAGAAACCCAGCTCAAGAGCGGAAAGCAA 2572
Qy 2522 CCGTCCCGAGCAGAACTGAGAGCTGGAAATAGAAATTAATTCCTCAACAGCTGTGCC 2581
Db 2573 CCGTCCCGAGCAGAACTGAGAGCTGGAAATAGAAATTAATTCCTCAACAGCTGTGCC 2632
Qy 2582 AGCCATTTTCTCTGGAGGAGCAGGAGCAATTTCTCTGCTGCTCAGCATCCAGAGCTCT 2641
Db 2633 AGCCATTTTCTCTGGAGGAGCAGGAGCAATTTCTCTGCTGCTCAGCATCCAGAGCTCT 2692
Qy 2642 CCTGTGCGATGACAGTGAAGAAACCCATCAAGGCTCTCAAAAGCTCGCGGACACCC 2701
Db 2693 CCTGTGCGATGACAGTGAAGAAACCCATCAAGGCTCTCAAAAGCTCGCGGACACCC 2752
Qy 2702 TGAGCTCAGGCGTACACTCTCTGGAGCAGCGAGGCTCGAAGCTCCAGCTGGAACA 2761
Db 2753 TGAGCTCAGGCGTACACTCTCTGGAGCAGCGAGGCTCGAAGCTCCAGCTGGAACA 2812

Qy 2762 TGGTCTGGCCCGGGGGCGGCCACCCGACACCCCAAGCTATTTCATTTGTTGTAAGTCC 2821
Db 2813 TGGTCTGGCCCGGGGGCGGCCACCCGACACCCCAAGCTATTTCATTTGTTGTAAGTCC 2872
Qy 2822 AATAACAGTCTCTTAATGTTGAACACCTGCAATCCGGGAGTTCCACCGGTTCAAAGTGG 2881
Db 2873 AATAACAGTCTCTTAATGTTGAACACCTGCAATCCGGGAGTTCCACCGGTTCAAAGTGG 2932
Qy 2882 GAGACATGCCACTGGCATCAGCAGCCAGATCCAGCTTCAGCTTGGTCAACA 2941
Db 2933 GAGACATGCCACTGGCATCAGCAGCCAGATCCAGCTTCAGCTTGGTCAACA 2992
Qy 2942 AAGACGGCAGCTCTGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGCAGT 3001
Db 2993 AAGACGGCAGCTCTGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGCAGT 3052
Qy 3002 GCACACTGGCCCCCTGATGGCAGCTTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGTGG 3061
Db 3053 GCACACTGGCCCCCTGATGGCAGCTTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGTGG 3112
Qy 3062 AGAACAGGCCCTAAACCTTGCCCTCCACCGCGGCTCCACATCGCCGGAAGCAGCTTCC 3121
Db 3113 AGAACAGGCCCTAAACCTTGCCCTCCACCGCGGCTCCACATCGCCGGAAGCAGCTTCC 3171
Qy 3122 TGTCTGGTGCAGATGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATGGCCAG 3181
Db 3172 TGTCTGGTGCAGATGCTGCTGCTG-AAAACAAGGCTCAGCGCTTCCAGGGATCTGCCAG 3230
Qy 3182 CCCCCGGCTCAAGTGGGAAACAGGGCTCCGACGACAGAGTGGGGGCAAGCAGAAATG 3241
Db 3231 -CCCCGGCTCAGCAGTGGGACAGGGCTCCGACGACAGAGTGGGGCAAGCAGAAATG 3289
Qy 3242 CTTCCAGGATTTACACCTGAGCCCTGCCACCCCTGCTGTAAGAAACA-TCGCGCAGCT 3300
Db 3290 CTTCCAGGATTTACACCTGAGCCCTGCCACCCCTGCTGAGAAAAACACTCCGCGCAGT 3349
Qy 3301 GAAGAGACAGAAAGAGGATGCGAGGTTTACTGGGGGAAAAACAAACAGGATCTTTTCT 3360
Db 3350 GAAGAGACAG-AGGAGGATGGCAGGATTAATCTCGGGAAACAAACAGG-ATCTTCTC 3404
Qy 3361 GCCCTGTCTCAGTTCAGTTCGCTTCAGCCCTGGATCAGTGACATTTGTTGGCAGAC 3420
Db 3405 TGCCCTGTCTCAGTTCAGTTCGCTTCAGCCCTGATCAGTGACATTTGTTGGCAGAC 3464
Qy 3421 AGGGAGACAGCTTCAGCTGGGTGAGAGGGGTGGGCGAGCCCTTCGGCCCTCACC 3480
Db 3465 AGGGAGACAGCTTCAGCTGGGTGAGAGGGGTGGGCGAGCCCTTCGGCCCTCACC 3524
Qy 3481 CTCCAGGCTGTGTGAGAGTGTCAAGTGTGTAAGGGGCCAAACTCAGGTTCAAGTGCAGAA 3540
Db 3525 CTCCAGGCTGTGTGAGAGTGTCAAGTGTGTAAGGGGCCAAACTCAGGTTCAAGTGCAGAA 3584
Qy 3541 CCAGTTCAGCAGGTATGCCCGCCCTAGGTTAAGGGGGCCCTCTAAACCCCTTCGCTGCG 3600
Db 3585 CCAGTTCAGCAGGTATGCCCGCCCTAGGTTAAGGGGGCCCTCTAAACCCCTTCGCTGCG 3644
Qy 3601 CTCACCTGGCAGCTCACCCCTTTTGGGTGTAGGGGAAAGATGCTGACCCCTGGGAAG 3660
Db 3645 CTCACCTGGCAGCTCACCCCTTTTGGGTGTAGGGGAAAGATGCTGACCCCTGGGAAG 3704
Qy 3661 GCTCCCTGGTAGAATACACCACTTTTCAGGTTGTTGCAACACAGGTTCTGAGTTGACC 3720
Db 3705 GCTCCCTGGTAGAATACACCACTTTTCAGGTTGTTGCAACACAGGTTCTGAGTTGACC 3764
Qy 3721 TCTGGTTCAGCCAAAGGACCAAGAGGTGTGTAGTGAAGTGGTTCTCAGTCCCAGACA 3780
Db 3765 TCTGGTTCAGCCAAAGGACCAAGAGGTGTGTAGTGAAGTGGTTCTCAGTCCCAGACA 3824
Qy 3781 TGTGCCCTTTGCTGCTGGCTACCACTTTTCCCCAGACAGGAGCCCGAGCCCTTCA 3840
Db 3825 TGTGCCCTTTGCTGCTGGCTACCACTTTTCCCCAGACAGGAGCCCGAGCCCTTCA 3884

Qy	3841	GGCCGAGCACTGCCCGACATCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGT	3900
Db	3885	GGCCGAGCACTGCCCGACACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGT	3944
Qy	3901	GATCAGGATATGCTGTGACAGAAACAGTCTGTGTGATGGAATGATCAGTGTCTTAAGGAAG	3960
Db	3945	GATCAGGATATGCTGTGACAGAAACAGTCTGTGTGATGGAATGATCAGTGTCTTAAGGAAG	4004
Qy	3961	CAGCAGAGAGAGACG-TCCGCGCCGCCAGACCCCACTATCAGTGTCCAGCGTGTCTGGTTC	4019
Db	4005	CAGCAGAGAGAGACGCTCGCGCGCCAGACCCCACTATCAGTGTCCAGCGTGTCTGGTTC	4064
Qy	4020	CCAGAGCACAGCTCAGCATCACACTGACATCACCTCGCTGCGCCCTCGCCGACAGAGGTA	4079
Db	4065	CCAGAGCACAGCTCAGCATCACACTGACATCACCTCGCTGCGCCCTCGCCGACAGAGGTA	4124
Qy	4080	CTGCCGACGGCACATTTGCACCTCTCATGACCTCAAAGCACTTTCATGCTGCCCTCTGGCA	4139
Db	4125	CTGCCGACGGCACATTTGCACCTCTCATGACCTCAAAGCACTTTCATGCTGCCCTCTGGCA	4184
Qy	4140	GGCAGGGCAGGGCAGTGACACTGTAGGAGCATGCAAGCAGAGATGGGTGAAGGGA	4199
Db	4185	GGCAGGGCAGGGCAGTGACACTGTAGGAGCATGCAAGCAGAGATGGGTGAAGGGA	4244
Qy	4200	CACAGTCTTGAGCTGTCCACATGATGTGACTCTCTCAAACCTCTTCACAGATTTCTTAAG	4259
Db	4245	CACAGTCTTGAGCTGTCCACATGATGTGACTCTCTCAAACCTCTTCACAGATTTCTTAAG	4304
Qy	4260	AATAGCACCCCTTCCCATTGGCCAGCTTAGCCTCTTCCAGGGGAGCTACTCAGG	4319
Db	4305	AATAGCACCCCTTCCCATTGGCCAGCTTAGCCTCTTCCAGGGGAGCTACTCAGG	4364
Qy	4320	ACTCACGTAGCATTAATCAGCTGTGAATCGTCAAGGGGTGTCTGTAGCGCTCAACCTCC	4379
Db	4365	ACTCACGTAGCATTAATCAGCTGTGAATCGTCAAGGGGTGTCTGTAGCGCTCAACCTCC	4424
Qy	4380	TGGGGCAGGGGACGCCGAGACTCGTGGGAGAGACTCATTTCCACACTTTTGGCAAGACAG	4439
Db	4425	TGGGGCAGGGGACGCCGAGACTCGTGGGAGAGACTCATTTCCACACTTTTGGCAAGACAG	4484
Qy	4440	CCTTTGTGCAGCTGTCCACATTCAGTCAGACTGCTCCCGGGGAGAGCGCCCGCCCCCA	4499
Db	4485	CCTTTGTGCAGCTGTCCACATTCAGTCAGACTGCTCCCGGGGAGAGCGCCCGCCCCCA	4544
Qy	4500	GCATATAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGGTGTGTAGAGAACTCTTTGTAA	4559
Db	4545	GCATATAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGGTGTGTAGAGAACTCTTTGTAA	4604
Qy	4560	GCAATAAAGTTTGGGGTGATGACAAATGTTT	4589
Db	4605	GCAATAAAGTTTGGGGTGATGACAAATGTTT	4634

Search completed: June 11, 2005, 04:01:20
Job time : 11508.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 09:44:05 ; Search time 1758.13 Seconds
(without alignments)
11471.453 Million cell updates/sec

Title: US-09-155-676B-6
Perfect score: 4596
Sequence: 1 agcgggggactgtgccgtg.....gatgacaaatgttaaaaaa 4596

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 12985247 seqs, 219411492 residues

Total number of hits satisfying chosen parameters: 25970494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA New:*

- 1: /cgn2_6/ptodata/2/pna/pct NEW COMB.seq:*
- 2: /cgn2_6/ptodata/2/pna/pct NEW COMB.seq:*
- 3: /cgn2_6/ptodata/2/pna/US06 NEW COMB.seq:*
- 4: /cgn2_6/ptodata/2/pna/US07 NEW COMB.seq:*
- 5: /cgn2_6/ptodata/2/pna/US08 NEW COMB.seq:*
- 6: /cgn2_6/ptodata/2/pna/US09 NEW COMB.seq:*
- 7: /cgn2_6/ptodata/2/pna/US10 NEW COMB.seq:*
- 8: /cgn2_6/ptodata/2/pna/US10 NEW COMB.seq:*
- 9: /cgn2_6/ptodata/2/pna/US10 NEW COMB.seq:*
- 10: /cgn2_6/ptodata/2/pna/US10 NEW COMB.seq:*
- 11: /cgn2_6/ptodata/2/pna/US11 NEW COMB.seq:*
- 12: /cgn2_6/ptodata/2/pna/US11 NEW COMB.seq:*
- 13: /cgn2_6/ptodata/2/pna/US11 NEW COMB.seq:*
- 14: /cgn2_6/ptodata/2/pna/US11 NEW COMB.seq:*
- 15: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4596	100.0	4596	6	US-09-155-676B-6
2	4578.4	99.6	4596	8	US-10-450-763-15294
3	4437.6	96.6	4684	15	US-60-659-397-566
4	4243.2	92.3	4458	15	US-60-659-397-567
5	1851.2	40.3	2631	6	US-09-155-676B-3
6	1554.2	33.8	65967	15	US-60-659-397-12129
7	1474.4	32.1	2760	8	US-10-450-763-15295
8	1345.6	29.3	1400	13	US-11-060-756-560
9	1345.6	29.3	1400	13	US-11-060-756-4832
10	676.2	14.7	3658	15	US-60-680-544-32424
11	676.2	14.7	3658	15	US-60-680-473-32424
12	627.6	13.7	722	8	US-10-450-763-15296
13	440.6	9.6	761	15	US-60-680-544-10785
14	440.6	9.6	761	15	US-60-680-544-11056
15	440.6	9.6	761	15	US-60-680-473-10785
16	440.6	9.6	761	15	US-60-680-473-11056
17	200.6	4.4	201	15	US-60-659-397-7882
18	200.6	4.4	201	15	US-60-659-397-7883

19	200.6	4.4	201	15	US-60-659-397-7885	Sequence 7885, Ap
20	200.6	4.4	201	15	US-60-659-397-7886	Sequence 7886, Ap
21	200.6	4.4	201	15	US-60-659-397-7887	Sequence 7887, Ap
22	200.6	4.4	201	15	US-60-659-397-7888	Sequence 7888, Ap
23	200.6	4.4	201	15	US-60-659-397-7889	Sequence 7889, Ap
24	200.6	4.4	201	15	US-60-659-397-7891	Sequence 7891, Ap
25	200.6	4.4	201	15	US-60-659-397-7892	Sequence 7892, Ap
26	200.6	4.4	201	15	US-60-659-397-7893	Sequence 7893, Ap
27	200.6	4.4	201	15	US-60-659-397-32722	Sequence 32722, A
28	200.6	4.4	201	15	US-60-659-397-32724	Sequence 32724, A
29	200.6	4.4	201	15	US-60-659-397-32758	Sequence 32758, A
30	153.2	3.3	201	15	US-60-659-397-7884	Sequence 7884, Ap
31	153.2	3.3	201	15	US-60-659-397-7890	Sequence 7890, Ap
32	153.2	3.3	201	15	US-60-659-397-32757	Sequence 32757, A
33	150	3.3	201	15	US-60-659-397-32751	Sequence 32751, A
34	136.8	3.0	493	1	PCT-US05-00517-813	Sequence 813, App
35	128.6	2.8	201	15	US-60-659-397-32755	Sequence 32755, A
36	119.2	2.6	1353	15	US-60-680-544-37753	Sequence 37753, A
37	119.2	2.6	1353	15	US-60-680-473-37753	Sequence 37753, A
38	98	2.1	201	15	US-60-659-397-32710	Sequence 32710, A
39	88.8	1.9	2207	11	US-10-940-774A-5071	Sequence 5071, Ap
40	88	1.9	201	15	US-60-659-397-32726	Sequence 32726, A
41	83.8	1.8	3089	6	US-09-608-890A-9	Sequence 9, Appli
42	78	1.7	201	15	US-60-659-397-32752	Sequence 32752, A
43	78	1.7	599	9	US-10-972-079-7876	Sequence 7876, Ap
44	75.2	1.6	600	9	US-10-972-079-7875	Sequence 7875, Ap
45	70	1.5	1024	8	US-10-450-763-15297	Sequence 15297, A

ALIGNMENTS

RESULT 1

US-09-155-676B-6
; Sequence 6, Application US/09155676B
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR
; FILE OF INVENTION: PREPARATION AND USE
; FILE REFERENCE: WALLACH-21
; CURRENT APPLICATION NUMBER: US/09/155.676B
; CURRENT FILING DATE: 1999-01-04
; PRIOR APPLICATION NUMBER: PCT/IL97/00117
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: IL 117800
; PRIOR FILING DATE: 1996-04-02
; PRIOR APPLICATION NUMBER: IL 119133
; PRIOR FILING DATE: 1996-08-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-155-676B-6

Query Match 100.0%; Score 4596; DB 6; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCGGGGGGACTGTGCCGTGTGGAACGTGTAGCTGTGTGAAGTGGACTCTGTATTACCATTTG 60

Db 1 AGCGGGGGGACTGTGCCGTGTGGAACGTGTAGCTGTGTGAAGTGGACTCTGTATTACCATTTG 60

Qy 61 AGAGTGTTCGAGGATGAGTATGTGTGGCAGAGCACATATAACAGGCAGAGACCCCTTT 120

Db 61 AGAGTGTTCGAGGATGAGTATGTGTGGCAGAGCACATATAACAGGCAGAGACCCCTTT 120

Qy 121 GCCCTGCGCTTTCTCCCAACCAAGGCTGACCTGTGTTCTCCCGAGTCTGGGATTCTA 180

|||||
121 GCCCTGCTTTCTCCCCCAACCCAAAGGCTGACCTGTGTTCTCCAGAGTCTGGATTCTA 180
QY
181 AGTGACCTGCTCTGTGTTGGTCTCTCTCAGATGAGCACAAGCCCTGGAGATGCGAGTG 240
Db
181 AGTGACCTGCTCTGTGTTGGTCTCTCTCAGATGAGCACAAGCCCTGGAGATGCGAGTG 240
QY
241 ATGGAAATGGCTGCTGCCAGTGCCTGCTCAGAGTGGGGCAGCAGAAGAACTCCCC 300
Db
241 ATGGAAATGGCTGCTGCCAGTGCCTGCTCAGAGTGGGGCAGCAGAAGAACTCCCC 300
QY
301 AAGCCAAAGGAGAGACGCGCCACTTGGGGAAAGAACAGAGTCCGTCTTACAAGTTGAG 360
Db
301 AAGCCAAAGGAGAGACGCGCCACTTGGGGAAAGAACAGAGTCCGTCTTACAAGTTGAG 360
QY
361 GCCGTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGACGTGATTACC 420
Db
361 GCCGTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGACGTGATTACC 420
QY
421 AAGGCGACAGCCAAAGAGGCTCCGAGCGAGGCGCAGCTGCCATCTCTATCATCGCCAG 480
Db
421 AAGGCGACAGCCAAAGAGGCTCCGAGCGAGGCGCAGCTGCCATCTCTATCATCGCCAG 480
QY
481 GCTGAGTGTGAAATAGCAAGAGTTTCAAGCCCACTTTTCAAGAACGCAATTTTTCATCGCT 540
Db
481 GCTGAGTGTGAAATAGCAAGAGTTTCAAGCCCACTTTTCAAGAACGCAATTTTTCATCGCT 540
QY
541 GGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCCAT 600
Db
541 GGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCCAT 600
QY
601 GCTACAGAGGGCAAAATGSCCGTGTGTTGGAAGGAAAGCGTCGAGCAAAAGCCCGG 660
Db
601 GCTACAGAGGGCAAAATGSCCGTGTGTTGGAAGGAAAGCGTCGAGCAAAAGCCCGG 660
QY
661 AAGAAACGGAAGAAAGAGCTCAAAGTCCCTGGCTCATGCAAGAGTGGCCCTTGGCCAAA 720
Db
661 AAGAAACGGAAGAAAGAGCTCAAAGTCCCTGGCTCATGCAAGAGTGGCCCTTGGCCAAA 720
QY
721 CCCCTCCAGAGCCCTGAGCAGAGAGTGCATCCAGTCCGCTCATGCAAGAGTGGCCCTTGGCCAAA 780
Db
721 CCCCTCCAGAGCCCTGAGCAGAGAGTGCATCCAGTCCGCTCATGCAAGAGTGGCCCTTGGCCAAA 780
QY
781 CCACCTCGGCGCCCATATGTTAGAAACACCCGCGAGTTTCAACAGCTCTGAAAGAACCA 840
Db
781 CCACCTCGGCGCCCATATGTTAGAAACACCCGCGAGTTTCAACAGCTCTGAAAGAACCA 840
QY
841 GGCCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGCCTACGGCCGGCTCTGCTCGA 900
Db
841 GGCCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGCCTACGGCCGGCTCTGCTCGA 900
QY
901 TCAGAACTCCAAACTGATGATGAGCCCTTGGCAATGTCTGAAACCAAGTGTGAAACTGCAAC 960
Db
901 TCAGAACTCCAAACTGATGATGAGCCCTTGGCAATGTCTGAAACCAAGTGTGAAACTGCAAC 960
QY
961 CACCCCGAGGAGGGCCCCCTGCCCCGCGCCAGCACCCTTCCCTATAGCAGACTG 1020
Db
961 CACCCCGAGGAGGGCCCCCTGCCCCGCGCCAGCACCCTTCCCTATAGCAGACTG 1020
QY
1021 CCTCATCTCCCTTCCCATTTCCACCTCTCCAGCCCTTGAAACCTTCAACCTCTGAGTCCCTTC 1080
Db
1021 CCTCATCTCCCTTCCCATTTCCACCTCTCCAGCCCTTGAAACCTTCAACCTCTGAGTCCCTTC 1080
QY
1081 CTGGGAAACTGGCTGTGTAGACAGCAGAAACCTTGGCTGTGACCCACACCTGTAGGCAAA 1140
Db
1081 CTGGGAAACTGGCTGTGTAGACAGCAGAAACCTTGGCTGTGACCCACACCTGTAGGCAAA 1140
QY
1141 CTGGCTGTGTAGACAGTCCAAAGCCCTGCTGCCCCACACCTTGAGCCAGCTGCTG 1200
Db
1141 CTGGCTGTGTAGACAGTCCAAAGCCCTGCTGCCCCACACCTTGAGCCAGCTGCTG 1200
QY
1201 TCTCGTGTGCCCATGAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGC 1260

Db
1201 TCTCGTGTGCCCATGAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGC 1260
QY
1261 AGCGTGAGCTCAAGCCAGGCCCAAGCTGACCTGAGCCCTGGCCCAAGACCTGGCAGCAGCG 1320
Db
1261 AGCGTGAGCTCAAGCCAGGCCCAAGCTGACCTGAGCCCTGGCCCAAGACCTGGCAGCAGCG 1320
QY
1321 GGCTCCAGATCCCGGGAGCCAGGCCCAAAACTTGAAGCAACAGAGGTGTCTGTCTCACT 1380
Db
1321 GGCTCCAGATCCCGGGAGCCAGGCCCAAAACTTGAAGCAACAGAGGTGTCTGTCTCACT 1380
QY
1381 GAGAACTCAAGCCAGTGGATTTATAGTATCCGAGAAGTCCACTGGGCCACGACCCAG 1440
Db
1381 GAGAACTCAAGCCAGTGGATTTATAGTATCCGAGAAGTCCACTGGGCCACGACCCAG 1440
QY
1441 CTCGCCCTGGCAGAGGCTCCTTGGAGAGTGCACAGATGGAGGACAAAGCAGACTGGC 1500
Db
1441 CTCGCCCTGGCAGAGGCTCCTTGGAGAGTGCACAGATGGAGGACAAAGCAGACTGGC 1500
QY
1501 TTCCAGTGCCTGTCAAAAAGGTGCGCTCGAAAGTATTTCCGGGCAAGAGAGCTGATGGCA 1560
Db
1501 TTCCAGTGCCTGTCAAAAAGGTGCGCTCGAAAGTATTTCCGGGCAAGAGAGCTGATGGCA 1560
QY
1561 TGTGAGGATTGACCTCAAGAGGATTTCTCCCTTTGTATGAGCTGTGAGAGAGGGCT 1620
Db
1561 TGTGAGGATTGACCTCAAGAGGATTTCTCCCTTTGTATGAGCTGTGAGAGAGGGCT 1620
QY
1621 TGGGTCAACATCTTCAATGAGCTGCTGAAAGTGGCTCCCTGGGCCAGCTGGTCAAGGAG 1680
Db
1621 TGGGTCAACATCTTCAATGAGCTGCTGAAAGTGGCTCCCTGGGCCAGCTGGTCAAGGAG 1680
QY
1681 CAGGCTGTCTCCACAGAGGACCGGCGCTGTACTACTCTGGGCCAGGCCCTGGAGGCTGTG 1740
Db
1681 CAGGCTGTCTCCACAGAGGACCGGCGCTGTACTACTCTGGGCCAGGCCCTGGAGGCTGTG 1740
QY
1741 GAATACCTCCACTCAAGAAAGATTCTGCATGGGAGAGCTCAAGCTGTGACAACTGTCTCTG 1800
Db
1741 GAATACCTCCACTCAAGAAAGATTCTGCATGGGAGAGCTCAAGCTGTGACAACTGTCTCTG 1800
QY
1801 TCCAGCGATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGCTTCAACCT 1860
Db
1801 TCCAGCGATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGCTTCAACCT 1860
QY
1861 GATGCCCTGGGAAAGTCTTCTCACAGGAGATACATCTCTGGCCACAGAGACCCACATG 1920
Db
1861 GATGCCCTGGGAAAGTCTTCTCACAGGAGATACATCTCTGGCCACAGAGACCCACATG 1920
QY
1921 GCTCCGAGGTGGTCTGGGCAAGAGTGTGCAACGCAAGGTGGATGTCTGAGAGAGCTGC 1980
Db
1921 GCTCCGAGGTGGTCTGGGCAAGAGTGTGCAACGCAAGGTGGATGTCTGAGAGAGCTGC 1980
QY
1981 TGTATGATGCTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTCCGAGGGCG 2040
Db
1981 TGTATGATGCTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTCCGAGGGCG 2040
QY
2041 CTCTGCCCTCAAGATTGGCCAGAGGCTCCGCTGTGAGGGAGATCCACCTCTCTGCGCC 2100
Db
2041 CTCTGCCCTCAAGATTGGCCAGAGGCTCCGCTGTGAGGGAGATCCACCTCTCTGCGCC 2100
QY
2101 CCTCTCACAGCCAGGCCATCAAGAGGGGCTGAGGAAAGAGCCCATCAACCGCTGTCT 2160
Db
2101 CCTCTCACAGCCAGGCCATCAAGAGGGGCTGAGGAAAGAGCCCATCAACCGCTGTCT 2160
QY
2161 GCAGGGAGCTGGGAGGAGGTTGAAACGGGCTGCCACCTGAGCTCAGTTCTGAGAGGCTGAGAGC 2220
Db
2161 GCAGGGAGCTGGGAGGAGGTTGAAACGGGCTGCCACCTGAGCTCAGTTCTGAGAGGCTGAGAGC 2220
QY
2221 CCTTGGAGGGAGGAATATAAAGAACCAAGACATCCACCCCAAAATCAAGCCAAATACCAC 2280
Db
2221 CCTTGGAGGGAGGAATATAAAGAACCAAGACATCCACCCCAAAATCAAGCCAAATACCAC 2280
QY
2281 CAGACCTTCCATGCCCAGCCAGAGAGCTTTGCGCAAGGGCCCCAGGGCCCCGGCAGCT 2340
Db
2281 CAGACCTTCCATGCCCAGCCAGAGAGCTTTGCGCAAGGGCCCCAGGGCCCCGGCAGCT 2340

Qy	2341	GAGGAGACAA	CAGGAGAGCC	CCCTAAGCT	CCAGGCTCT	CTCTCCCA	CAGAGCC	CCCCAGAG	2400
Db	2341	GAGGAGACAA	CAGGAGAGCC	CCCTAAGCT	CCAGGCTCT	CTCTCCCA	CAGAGCC	CCCCAGAG	2400
Qy	2401	CCAAACAGT	CTCTCCCTG	CACTTTGAG	CAAGGAGT	CTGGAT	TCTGGAT	TCTGGAT	2460
Db	2401	CCAAACAGT	CTCTCCCTG	CACTTTGAG	CAAGGAGT	CTGGAT	TCTGGAT	TCTGGAT	2460
Qy	2461	CCTCTGCT	CCCTGGAG	CCAGCCCT	CGCAGAA	CCCCAGCT	CACAGAG	CGGAAAGCA	2520
Db	2461	CCTCTGCT	CCCTGGAG	CCAGCCCT	CGCAGAA	CCCCAGCT	CACAGAG	CGGAAAGCA	2520
Qy	2521	ACCTGCCG	AGCAGGAA	CTCGAG	CTGGAAT	AGAAAT	TATTCCT	CAACAG	2580
Db	2521	ACCTGCCG	AGCAGGAA	CTCGAG	CTGGAAT	AGAAAT	TATTCCT	CAACAG	2580
Qy	2581	CAGCCAT	TTTTCTG	GAGGAG	CAGGCAAA	TTCTCTG	TGCTC	CAGCAT	2640
Db	2581	CAGCCAT	TTTTCTG	GAGGAG	CAGGCAAA	TTCTCTG	TGCTC	CAGCAT	2640
Qy	2641	TCCCTGCG	ATGATG	AGTGAAG	AAACCA	CTCAAG	GCTCG	CGGACAC	2700
Db	2641	TCCCTGCG	ATGATG	AGTGAAG	AAACCA	CTCAAG	GCTCG	CGGACAC	2700
Qy	2701	CTGAGCT	CAGGCGT	ACACTCT	CTGAGC	AGCGCG	AGGCTCG	AAAGCT	2760
Db	2701	CTGAGCT	CAGGCGT	ACACTCT	CTGAGC	AGCGCG	AGGCTCG	AAAGCT	2760
Qy	2761	ATGTGTCT	GCGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	2820
Db	2761	ATGTGTCT	GCGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	2820
Qy	2821	CAAAATAG	CTCTTTA	TGTTG	TAACCTG	CAACAT	CGGGAG	TCCAC	2880
Db	2821	CAAAATAG	CTCTTTA	TGTTG	TAACCTG	CAACAT	CGGGAG	TCCAC	2880
Qy	2881	GGAGACAT	TCGCAC	TGAGCAG	AGCCAG	ATCCAG	CTGAG	CTTGGT	2940
Db	2881	GGAGACAT	TCGCAC	TGAGCAG	AGCCAG	ATCCAG	CTGAG	CTTGGT	2940
Qy	2941	AAAGCGG	CAGCCT	GTTCGCT	AGCATG	AGAGT	CGCAG	ACTCGG	3000
Db	2941	AAAGCGG	CAGCCT	GTTCGCT	AGCATG	AGAGT	CGCAG	ACTCGG	3000
Qy	3001	TGCACTG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	3060
Db	3001	TGCACTG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	3060
Qy	3061	GAGAACAG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	3120
Db	3061	GAGAACAG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	3120
Qy	3121	CTGCTCG	GTGCA	CGATGCT	CGGCGG	CGGCGG	CGGCGG	CGGCGG	3180
Db	3121	CTGCTCG	GTGCA	CGATGCT	CGGCGG	CGGCGG	CGGCGG	CGGCGG	3180
Qy	3181	GGCCCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	3240
Db	3181	GGCCCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	3240
Qy	3241	GCCTCCAG	GATTTT	CACTG	AGCCCT	CGCAG	CAAGT	TCCGCC	3300
Db	3241	GCCTCCAG	GATTTT	CACTG	AGCCCT	CGCAG	CAAGT	TCCGCC	3300
Qy	3301	GAAGAGA	CAGAGG	AGTAC	CTGGG	AAACAA	CAAGG	ATCTTT	3360
Db	3301	GAAGAGA	CAGAGG	AGTAC	CTGGG	AAACAA	CAAGG	ATCTTT	3360
Qy	3361	GGCCCGT	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	3420
Db	3361	GGCCCGT	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	3420

Qy	3421	AGGGAGAG	CAGGCTT	CCAGGCT	CGGCTC	AGAGGG	GGGCGG	CGGCTT	CGGCGG	3480
Db	3421	AGGGAGAG	CAGGCTT	CCAGGCT	CGGCTC	AGAGGG	GGGCGG	CGGCTT	CGGCGG	3480
Qy	3481	CTCCAGG	CTGCTG	AGAGTGT	CAAGTGT	TAAGGG	CGCCAA	CTCAGG	TTCAAGT	3540
Db	3481	CTCCAGG	CTGCTG	AGAGTGT	CAAGTGT	TAAGGG	CGCCAA	CTCAGG	TTCAAGT	3540
Qy	3541	CCAGGT	CAGCAG	GTATG	CCCGCG	GTAGGTT	TAAGGG	GGGCGG	CGCTT	3600
Db	3541	CCAGGT	CAGCAG	GTATG	CCCGCG	GTAGGTT	TAAGGG	GGGCGG	CGCTT	3600
Qy	3601	CTCAGCT	GGCAGG	CTCAG	CCCTTT	TGCGT	AGGAA	AGATG	CCCTG	3660
Db	3601	CTCAGCT	GGCAGG	CTCAG	CCCTTT	TGCGT	AGGAA	AGATG	CCCTG	3660
Qy	3661	GCTCCCT	GTTAG	AAATAC	CAACAC	CTTTT	CAAGT	TTTGT	CAACAC	3720
Db	3661	GCTCCCT	GTTAG	AAATAC	CAACAC	CTTTT	CAAGT	TTTGT	CAACAC	3720
Qy	3721	TCTGTTT	CAGC	CAAGG	AGTGT	TAAGT	GAAGT	TGTTT	CTCAGT	3780
Db	3721	TCTGTTT	CAGC	CAAGG	AGTGT	TAAGT	GAAGT	TGTTT	CTCAGT	3780
Qy	3781	TGTGCCC	CTTTG	CTGCT	GGCTT	ACCACT	CTTTT	CCAG	AGCAG	3840
Db	3781	TGTGCCC	CTTTG	CTGCT	GGCTT	ACCACT	CTTTT	CCAG	AGCAG	3840
Qy	3841	GGCCCA	GACTG	CCCCAG	ACTG	CGTGG	CACTG	CGTGG	CACTG	3900
Db	3841	GGCCCA	GACTG	CCCCAG	ACTG	CGTGG	CACTG	CGTGG	CACTG	3900
Qy	3901	GATG	CAGG	ATGCT	CTG	CAAGG	AACTG	CTG	TGATG	3960
Db	3901	GATG	CAGG	ATGCT	CTG	CAAGG	AACTG	CTG	TGATG	3960
Qy	3961	CAGCAG	AGAG	AGCTG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	4020
Db	3961	CAGCAG	AGAG	AGCTG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	CGGCGG	4020
Qy	4021	CAGAGC	ACAG	CTCAG	CACTG	CACTG	CACTG	CACTG	CACTG	4080
Db	4021	CAGAGC	ACAG	CTCAG	CACTG	CACTG	CACTG	CACTG	CACTG	4080
Qy	4081	TGCCAG	CGGCA	CTTGT	GACTT	CTGAT	GACTT	CTGAT	GACTT	4140
Db	4081	TGCCAG	CGGCA	CTTGT	GACTT	CTGAT	GACTT	CTGAT	GACTT	4140
Qy	4141	GGCAGG	CGGCG	AGTGA	CACTG	TAGG	AGCAT	AGCAAG	CGGAG	4200
Db	4141	GGCAGG	CGGCG	AGTGA	CACTG	TAGG	AGCAT	AGCAAG	CGGAG	4200
Qy	4201	ACAGT	CTT	GAGT	GTCC	CAATG	GTCA	CTCT	CAAA	4260
Db	4201	ACAGT	CTT	GAGT	GTCC	CAATG	GTCA	CTCT	CAAA	4260
Qy	4261	ATAGC	ACCC	CTTCC	CAATG	GTCC	CAATG	GTCC	CAATG	4320
Db	4261	ATAGC	ACCC	CTTCC	CAATG	GTCC	CAATG	GTCC	CAATG	4320
Qy	4321	CTCAGT	AGAT	TTAAAT	CAGT	GTGA	ATCG	GTG	AGG	4380
Db	4321	CTCAGT	AGAT	TTAAAT	CAGT	GTGA	ATCG	GTG	AGG	4380
Qy	4381	GGGCG	AGG	AGCG	CGG	AGAT	CTCG	TGG	AGAG	4440
Db	4381	GGGCG	AGG	AGCG	CGG	AGAT	CTCG	TGG	AGAG	4440
Qy	4441	CTTTGT	CCAG	GTG	CCAC	TTT	GAGT	CTG	CTCC	4500
Db	4441	CTTTGT	CCAG	GTG	CCAC	TTT	GAGT	CTG	CTCC	4500
Qy	4501	CACA	TAAG	AAAT	CTG	CAG	AGT	CTG	GGG	4560

Db 4501 CACATAAAGAACTGCAGCCCTTGGTACTGCAGAGTCTGGGTCTGAGAGACCTTTGTGAAG 4560
QY 4561 CAATAAAGTTTGGGTGATGACAAATGTTAAAAAAA 4596
Db 4561 CAATAAAGTTTGGGTGATGACAAATGTTAAAAAAA 4596

RESULT 2
US-10-450-763-15294
; Sequence 15294, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 15294
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1117)..(1581)
; OTHER INFORMATION: 100% homologous to Homo sapiens sperm protein 10, accession
; OTHER INFORMATION: number M82968, Smith-Waterman Score=809.
US-10-450-763-15294

Query Match 99.6%; Score 4578.4; DB 8; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 5 GGGGGACTGTGCGTGTGAAACGTGTAGCTGTGA-AGGTGAATCTGTACCATTGAGG 63
Db 5 GGGGGACTGTGCGTGTGAAACGTGTAGCTGTGAAGGTGACTCTGTACCATTGAGG 64

QY 64 ATGTTTGGAGTAGTAGTGTGTCAGAGGCACACATAAAGCAGCAGACCCCTTGGC 123
Db 65 ATGTTTGGAGTAGTAGTGTGTCAGAGGCACACATAAAGCAGCAGACCCCTTGGC 124

QY 124 CTGCTCTTTCTCCCCAACCCCAAGGCTGACCTGTGTCTCCAGAGTCTGGGATTTCTAAGT 183
Db 125 CTGCTCTTTCTCCCCAACCCCAAGGCTGACCTGTGTCTCCAGAGTCTGGGATTTCTAAGT 184

QY 184 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCAACAAGCTGGAGATGGCAATGATG 243
Db 185 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCAACAAGCTGGAGATGGCAATGATG 244

QY 244 GAAATGGCTGCCAGGTCGCCCTGGCTCAGCAGTGGGAGCAAGAACTCTCCCAAG 303
Db 245 GAAATGGCTGCCAGGTCGCCCTGGCTCAGCAGTGGGAGCAAGAACTCTCCCAAG 304

QY 304 CCAAGAGGAGAACGCGCCCACTGGGAGAAACACAGAGCTCCGTCTTACAAGCTTGAGGCC 363
Db 305 CCAAGAGGAGAACGCGCCCACTGGGAGAAACACAGAGCTCCGTCTTACAAGCTTGAGGCC 364

QY 364 GTGAGAGAGCCCTGTGTTCTTCGGAAGAGTGGAGATCCTGAATGAAGTGAATTAACAAG 423
Db 365 GTGAGAGAGCCCTGTGTTCTTCGGAAGAGTGGAGATCCTGAATGAAGTGAATTAACAAG 424

QY 424 GGCACAGCCAAAGAGGCTCCAGGAGGCGCAGCTGCCATCTCTATCATCCCGCAGGCT 483
Db 425 GGCACAGCCAAAGAGGCTCCAGGAGGCGCAGCTGCCATCTCTATCATCCCGCAGGCT 484

QY 484 GAGTGTGAGAAATAGCCAAAGATTTCAGCCCCACCTTTTCAGAACGCAATTTTCATCGCTGGG 543

Db 485 GAGTGTGAGAAATAGCCAAAGATTTCAGCCCCACCTTTTCAGAACGCAATTTTCATCGCTGGG 544
QY 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCAACAAATGTGGGCCCATGCT 603
Db 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCAACAAATGTGGGCCCATGCT 604

QY 604 ACAGAGGGCAAAATGGCCCGCTGTGTTGGAGGAAAGCGTCGAGCAAAAGCCCGGAAG 663
Db 605 ACAGAGGGCAAAATGGCCCGCTGTGTTGGAGGAAAGCGTCGAGCAAAAGCCCGGAAG 664

QY 664 AAAACGAAAGAAAGAGAGCTCAAAGTCCCTGGCTCATGCAAGAGTGGCCTTGCCCAAAACC 723
Db 665 AAAACGAAAGAAAGAGAGCTCAAAGTCCCTGGCTCATGCAAGAGTGGCCTTGCCCAAAACC 724

QY 724 CTCCCCAGGACCCCTGTGAGCAGAGAGTGCACCAATCCCAAGTGCAGGAGGATGAGTCTCA 783
Db 725 CTCCCCAGGACCCCTGTGAGCAGAGAGTGCACCAATCCCAAGTGCAGGAGGATGAGTCTCA 784

QY 784 CTGGCGGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCCTCTCGAAGAACCCAGGC 843
Db 785 CTGGCGGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCCTCTCGAAGAACCCAGGC 844

QY 844 CTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTTACGCGCGGCTCTGCTCGATCA 903
Db 845 CTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTTACGCGCGGCTCTGCTCGATCA 904

QY 904 GAACTCCACAAACTGATCAGGCCCTTGTGCAATGTCTGAACCAAGTGTGAAACTGCAACAC 963
Db 905 GAACTCCACAAACTGATCAGGCCCTTGTGCAATGTCTGAACCAAGTGTGAAACTGCAACAC 964

QY 964 CCCAGGAGGAGGCGCCCTGCGCCCGCCACGACCCCTTCCCTATGAGAGACTGCT 1023
Db 965 CCCAGGAGGAGGCGCCCTGCGCCCGCCACGACCCCTTCCCTATGAGAGACTGCT 1024

QY 1024 CATCCCTTCCCAATTCACCCCTCTCCAGCCCTTGGAAACCTCACCTCTCTGGAGTCTTCTCTG 1083
Db 1025 CATCCCTTCCCAATTCACCCCTCTCCAGCCCTTGGAAACCTCACCTCTCTGGAGTCTTCTCTG 1084

QY 1084 GGCMAACTGGCCTGTGTAGACAGCAGAAACCTTGGCCTGACCCACACCTGAGCAAACTG 1143
Db 1085 GGCMAACTGGCCTGTGTAGACAGCAGAAACCTTGGCCTGACCCACACCTGAGCAAACTG 1144

QY 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTTGGAGCCAGCTGCTGCT 1203
Db 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTTGGAGCCAGCTGCTGCT 1204

QY 1204 CGTGTGCTCCATGAGAAAGTCTTGTGTGAGAAATACCTAGTGCATGCTCTGCAAGGCAGC 1263
Db 1205 CGTGTGCTCCATGAGAAAGTCTTGTGTGAGAAATACCTAGTGCATGCTCTGCAAGGCAGC 1264

QY 1264 GTGAGCTCAAGCCAGGCCCCACAGCTGACAGCTGGCCCAAGACCTGGGCGACACGGGGC 1323
Db 1265 GTGAGCTCAAGCCAGGCCCCACAGCTGACAGCTGGCCCAAGACCTGGGCGACACGGGGC 1324

QY 1324 TCCAGATCCCGGAGCCAGCCCAAAACTCAGGAAACACGAGGGTGTCTGCTCACTGAG 1383
Db 1325 TCCAGATCCCGGAGCCAGCCCAAAACTCAGGAAACACGAGGGTGTCTGCTCACTGAG 1384

QY 1384 AAATCAAGCCAGTGGATTATGAGTACCGAGAAGTCCACTGGGCCACACGACCAAGCTC 1443
Db 1385 AAATCAAGCCAGTGGATTATGAGTACCGAGAAGTCCACTGGGCCACACGACCAAGCTC 1444

QY 1444 GCCTGGGCGAGAGGCTCCTTCGAGAGGTGCACAGGATGGAGGACAAAGCAAGCTGGCTTC 1503
Db 1445 GCCTGGGCGAGAGGCTCCTTCGAGAGGTGCACAGGATGGAGGACAAAGCAAGCTGGCTTC 1504

QY 1504 CAGTGCCTGTCAAAAGAGGTGCGCTGGAAGTATTTTCGGGCGAGAGAGCTGATGCAATGT 1563
Db 1505 CAGTGCCTGTCAAAAGAGGTGCGCTGGAAGTATTTTCGGGCGAGAGAGCTGATGCAATGT 1564

QY 1564 GCAGGATTGACCTCAACCAGAAATTGCTCCTTTGTATGAGCTGTGAGAGAGGGCTTGG 1623

Db	1565	CGAGATTGACCTCACCCAGAAATTGTCCCTTTGTATGAGAGCTGTGAGAGAGGCGCTTGG	1624
Qy	1624	GTCAACATCTTTATGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAG	1683
Db	1625	GTCAACATCTTTATGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAG	1684
Qy	1684	GGCTGTCTCCAGAGGACCGGGCCCTGTATCTACCTTGGGCCAGGCCCTTGAGAGGCTCTGGAA	1743
Db	1685	GGCTGTCTCCAGAGGACCGGGCCCTGTATCTACCTTGGGCCAGGCCCTTGAGAGGCTCTGGAA	1744
Qy	1744	TACCTCCACTCACAAGGATTTCTCATGGGAGCTCAAAGCTGACAAGCTGCTCCTGTCC	1803
Db	1745	TACCTCCACTCACAAGGATTTCTCATGGGAGCTCAAAGCTGACAAGCTGCTCCTGTCC	1804
Qy	1804	AGCATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGTGTGTCTTTCAACCTGAT	1863
Db	1805	AGCATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGTGTGTCTTTCAACCTGAT	1864
Qy	1864	GGCTGGGAAAGTCTTGTCTCAGAGGGACTATCATCTCCCTGGCA CAGAGACCACATGGCT	1923
Db	1865	GGCTGGGAAAGTCTTGTCTCAGAGGGACTATCATCTCCCTGGCA CAGAGACCACATGGCT	1924
Qy	1924	CCGAGGTGGTGTGGGAGGAGCTGCGACCCCAAGGTGGATGTCTGGAGCAGCTGTGT	1983
Db	1925	CCGAGGTGGTGTGGGAGGAGCTGCGACCCCAAGGTGGATGTCTGGAGCAGCTGTGT	1984
Qy	1984	ATGATGTGACATGTCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTC	2043
Db	1985	ATGATGTGACATGTCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTC	2044
Qy	2044	TGCTCAAGATTTGCCAGGAGCTCCGCTGTGAGGAGGATCCACCCCTCTGTGGCCCT	2103
Db	2045	TGCTCAAGATTTGCCAGGAGCTCCGCTGTGAGGAGGATCCACCCCTCTGTGGCCCT	2104
Qy	2104	CTCAGCCAGGCCATCTCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA	2163
Db	2105	CTCAGCCAGGCCATCTCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA	2164
Qy	2164	CGGAGCTGGAGGGAAGGTGAACCGGSCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT	2223
Db	2165	CGGAGCTGGAGGGAAGGTGAACCGGSCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT	2224
Qy	2224	TGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAACAG	2283
Db	2225	TGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAACAG	2284
Qy	2284	ACCTTCATGCCAGCCAGAGAGCTTTGGCCAAAGGGCCCGAGGCCCGGCCAGCTGAG	2343
Db	2285	ACCTTCATGCCAGCCAGAGAGCTTTGGCCAAAGGGCCCGAGGCCCGGCCAGCTGAG	2344
Qy	2344	GAGACAAGCAGAGCCCTTAAGCTCCAGCTCCCTCTCCACACAGAGCCCGCCAGAGCCA	2403
Db	2345	GAGACAAGCAGAGCCCTTAAGCTCCAGCTCCCTCTCCACACAGAGCCCGCCAGAGCCA	2404
Qy	2404	AACAAGTCTCCTCCCTTGACTTTGAGCAAGAGGAGTCTGGGATGTGGGAAACCCCTTACCT	2463
Db	2405	AACAAGTCTCCTCCCTTGACTTTGAGCAAGAGGAGTCTGGGATGTGGGAAACCCCTTACCT	2464
Qy	2464	CTGTCTCCTGGAGCCAGCCCTTGCAGAAAACCCAGCTCACAGAGCGGAAAGCAACC	2523
Db	2465	CTGTCTCCTGGAGCCAGCCCTTGCAGAAAACCCAGCTCACAGAGCGGAAAGCAACC	2524
Qy	2524	GTCCGGAGCAGGAACCTGACAGCTGGAATAGAAATTTCTCAACAGCCTGTCCAG	2583
Db	2525	GTCCGGAGCAGGAACCTGACAGCTGGAATAGAAATTTCTCAACAGCCTGTCCAG	2584
Qy	2584	CCATTTCTCTGGAGGAGCAGGACAAATCTCTGTGCTCAGCATCGACAGCTCTCC	2643
Db	2585	CCATTTCTCTGGAGGAGCAGGACAAATCTCTGTGCTCAGCATCGACAGCTCTCC	2644
Qy	2644	CTGTCCGATGACAGTGAGAGAAACCCATCAAGAGCCCTCTCAAGCTCCGCGGACACCCCTG	2703
Db	2645	CTGTCCGATGACAGTGAGAGAAACCCATCAAGAGCCCTCTCAAGCTCCGCGGACACCCCTG	2704

Qy	2704	AGCTCAGCGGTACATCTCTGGAGCAGCCAGCCGAGGCTCGAAGCTCCAGCTGGAAACATG	2763
Db	2705	AGCTCAGCGGTACATCTCTGGAGCAGCCAGCCGAGGCTCGAAGCTCCAGCTGGAAACATG	2764
Qy	2764	GTCTGGCCCCGGGGGGCCACCGACACCCCAAGCTATTTCATGTGTGAAAGTCCAA	2823
Db	2765	GTCTGGCCCCGGGGGGCCACCGACACCCCAAGCTATTTCATGTGTGAAAGTCCAA	2824
Qy	2824	ATACAGTCTCTTAATGTGAAACACCTGCA CATCCGGAGTTCCACCGGTTCAAAAGTGGGA	2883
Db	2825	ATACAGTCTCTTAATGTGAAACACCTGCA CATCCGGAGTTCCACCGGTTCAAAAGTGGGA	2884
Qy	2884	GACATGCCACTGCGCATCAGCAGCAGATCCAGCTCGACCTTTCAGCTTGGTCAACAAA	2943
Db	2885	GACATGCCACTGCGCATCAGCAGCAGATCCAGCTCGACCTTTCAGCTTGGTCAACAAA	2944
Qy	2944	GACGGGAGCCTGTTCCGTACGATGGAGGTGCCAGCTCGGGCATCGACCTGCAGTGC	3003
Db	2945	GACGGGAGCCTGTTCCGTACGATGGAGGTGCCAGCTCGGGCATCGACCTGCAGTGC	3004
Qy	3004	ACACTGGCCCTGATGCGAGCTTCGCTGAGCTGAGGGTCAAGCATGGCCAGCTGGAG	3063
Db	3005	ACACTGGCCCTGATGCGAGCTTCGCTGAGCTGAGGGTCAAGCATGGCCAGCTGGAG	3064
Qy	3064	AACAGGCCCTAACCTGSCCTTCCACCGCCGCTCCACACTGCCGGAAGCAGCCTTCTCTG	3123
Db	3065	AACAGGCCCTAACCTGSCCTTCCACCGCCGCTCCACACTGCCGGAAGCAGCCTTCTCTG	3124
Qy	3124	CTCGTGCAGATGCTGCCCTGAAAAACAAGGCTCAGCGTTCCAGGGGATTTCCAGGCC	3183
Db	3125	CTCGTGCAGATGCTGCCCTGAAAAACAAGGCTCAGCGTTCCAGGGGATTTCCAGGCC	3184
Qy	3184	CCCGGCTCAGTGGGAAACAGGGCTCCAGCAGCAAGTGGGGCAGAGCAAGTGC	3243
Db	3185	CCCGGCTCAGTGGGAAACAGGGCTCCAGCAGCAAGTGGGGCAGAGCAAGTGC	3244
Qy	3244	TCCAGGATTTACACCTGAGCCCTGCCCCACCTGCTGAAAAAACAATCCGCCACGTGAA	3303
Db	3245	TCCAGGATTTACACCTGAGCCCTGCCCCACCTGCTGAAAAAACAATCCGCCACGTGAA	3304
Qy	3304	GAGACAGAGGAGATGGCAGGAGTTACCTGGGGAACAACAGGGATCTTTTTCGTC	3363
Db	3305	GAGACAGAGGAGATGGCAGGAGTTACCTGGGGAACAACAGGGATCTTTTTCGTC	3364
Qy	3364	CCTGCTCCAGTGGGCTGACCCGCTTGGATCAGTGACCATTTGTTGGCAGAGCAGG	3423
Db	3365	CCTGCTCCAGTGGGCTGACCCGCTTGGATCAGTGACCATTTGTTGGCAGAGCAGG	3424
Qy	3424	GGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGCGAGCCCTTCGGGCCCTCACCCCTC	3483
Db	3425	GGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGCGAGCCCTTCGGGCCCTCACCCCTC	3484
Qy	3484	CAGGTGCTGTGAGAGTCAAGTGTGTAAGGGCCAAAACCTCAGTTTCAAGTGCAGAACCA	3543
Db	3485	CAGGTGCTGTGAGAGTCAAGTGTGTAAGGGCCAAAACCTCAGTTTCAAGTGCAGAACCA	3544
Qy	3544	GGTTCAGAGTATGCCCGCGGTAGTTAAGGGGCCCTCTTAACCCCTTCGCTGGCCTC	3603
Db	3545	GGTTCAGAGTATGCCCGCGGTAGTTAAGGGGCCCTCTTAACCCCTTCGCTGGCCTC	3604
Qy	3604	ACCTGGCCAGCTCACCCCTTTGGGTGAGGGGAAAGAAATGCTGACCTTGGGAAAGGCT	3663
Db	3605	ACCTGGCCAGCTCACCCCTTTGGGTGAGGGGAAAGAAATGCTGACCTTGGGAAAGGCT	3664
Qy	3664	CCCTGGTAGAATACACACATTTTTCAGTTTGTGCAACA CAGGTCTCTGAGTTGACCTCT	3723
Db	3665	CCCTGGTAGAATACACACATTTTTCAGTTTGTGCAACA CAGGTCTCTGAGTTGACCTCT	3724
Qy	3724	GGTTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGTCTCAGTCCCCCAGACATGT	3783
Db	3725	GGTTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGTCTCAGTCCCCCAGACATGT	3784

Db 3172 TGCTCGGTGACGATGCTGCTGCCCTG-AAACACAGGCTCAGCCGTTTCACAGGGATCTGCCAG 3230
Qy 3182 CCCCCCGGCTCAGTGGGAACAGGCTCGCAGCAGCAAGTGGGGCAGAGCAATG 3241
Db 3231 -CCCCCGGCTCAGAGTGGGACAGGCTCGCAGCAGCAAGTGGGGCAGAGCAATG 3289
Qy 3242 CTTCCAGGATTCACACTGAGCCCTGCCCCACCCCTGCTGAAAAACA-TCGGCCAGT 3300
Db 3290 CTTCCAGGATTCACACTGAGCCCTGCCCCACCCCTGCTGAAAAACA-TCGGCCAGT 3349
Qy 3301 GAAGAGACAGAGGAGTGGCAGAGTTACCTGGGGAAAAAACAACAGGATCTTTTCT 3360
Db 3350 GAAGAGACAG-AGGAGGATGGCAGAGTTACCTGGGGAAAAAACAACAGG-ATCTTCTC 3404
Qy 3361 GCCCTGCTCAGTTCAGTGGCTTGACCCGCTTGATCAGTGACATTTGTTGCGAGAC 3420
Db 3405 TGCCCTGCTCAGTTCAGTGGCTTGACCCGCTTGATCAGTGACATTTGTTGCGAGAC 3464
Qy 3421 AGGGAGAGCAGCTTTCCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTGGGCCCTCACC 3480
Db 3465 AGGGAGAGCAGCTTTCCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTGGGCCCTCACC 3524
Qy 3481 CTCAGGCTGCTGTCAGAGTTCAGTGTGAAGGGCCCAACTCAGGTTAGTGCAGAA 3540
Db 3525 CTCAGGCTGCTGTCAGAGTTCAGTGTGAAGGGCCCAACTCAGGTTAGTGCAGAA 3584
Qy 3541 CCAGGTTCAGCAGGTATGCCCGCCCTGAGGTAAAGGGGGCCCTCTAAACCCCTTGCTGCGC 3600
Db 3585 CCAGGTTCAGCAGGTATGCCCGCCCTGAGGTAAAGGGGGCCCTCTAAACCCCTTGCTGCGC 3644
Qy 3601 CTCACCTGSCCAGCTCACCCCTTTTGGGTGTAAGGGGAAAGAAATGCTGACCTCGGAG 3660
Db 3645 CTCACCTGSCCAGCTCACCCCTTTTGGGTGTAAGGGGAAAGAAATGCTGACCTCGGAG 3704
Qy 3661 GTCCTGTTAGAAATACACCACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACC 3720
Db 3705 GCTCCTGTTAGAAATACACCACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACC 3764
Qy 3721 TCTGTTTCAGCAAGGACCAAGAGGTGTGAAGTGAAGTGTCTCAGTCCCCAGACA 3780
Db 3765 TCTGTTTCAGCAAGGACCAAGAGGTGTGAAGTGAAGTGTCTCAGTCCCCAGACA 3824
Qy 3781 TGTGCCCTTTGCTGCTGCTACCACTTTTCCAGAGCAGAGCCCGAGCCCTTTCA 3840
Db 3825 TGTGCCCTTTGCTGCTGCTACCACTTTTCCAGAGCAGAGCCCGAGCCCTTTCA 3884
Qy 3841 GSCCAGACCTGCCCCAGACTGCTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGT 3900
Db 3885 GSCCAGACCTGCCCCAGACTGCTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGT 3944
Qy 3901 GATCAGGATATGCTGACAGAAACAGTCTGTGGATGGACATGATCAGTGTCTAAGGAAAG 3960
Db 3945 GATCAGGATATGCTGACAGAAACAGTCTGTGGATGGACATGATCAGTGTCTAAGGAAAG 4004
Qy 3961 CAGCAGAGAGAGC-TCGGGGCCCCAGCCCCCATATCAGTGTCCAGCGTCTGTTCC 4019
Db 4005 CAGCAGAGAGAGCCTCGGGGGCCCCAGCCCCCATATCAGTGTCCAGCGTCTGTTCC 4064
Qy 4020 CCAGAGCAGCTCAGCATCACTGACACTCACTCACTGCGCTGCGCTGGCCAGAGGTA 4079
Db 4065 CCAGAGCAGCTCAGCATCACTGACACTCACTCACTGCGCTGCGCTGGCCAGAGGTA 4124
Qy 4080 CTGCGAGGCACTTTGCACTCTGATGACTCAAGACATTTTCATGGCTGCCCTTGCCA 4139
Db 4125 CTGCGAGGCACTTTGCACTCTGATGACTCAAGACATTTTCATGGCTGCCCTTGCCA 4184
Qy 4140 GGGCAGGCGAGGCGATGACACTGTAGGAGCATAGCAAGCCAGGAGTGGGTGAAGGA 4199
Db 4185 GGGCAGGCGAGGCGATGACACTGTAGGAGCATAGCAAGCCAGGAGTGGGTGAAGGA 4244
Qy 4200 CACAGTCTTGAAGTTCACATGCACTGATGACTCTCTCAAACTCTTCCAGATTTCTTAAG 4259
Db 4245 CACAGTCTTGAAGTTCACATGCACTGATGACTCTCTCAAACTCTTCCAGATTTCTTAAG 4304

Qy 4260 AATAGCAGCCCTTCCCATTTGCCAGCTTAGCCTTTCTTCCCAGGGAGCTACTCAGG 4319
Db 4305 AATAGCAGCCCTTCCCATTTGCCAGCTTAGCCTTTCTTCCCAGGGAGCTACTCAGG 4364
Qy 4320 ACTCAGTAGCATTAATCAGCTGTGAATCAGTGTGAGGGGTGTCTGCTAGCCTCAACCTCC 4379
Db 4365 ACTCAGTAGCATTAATCAGCTGTGAATCAGTGTGAGGGGTGTCTGCTAGCCTCAACCTCC 4424
Qy 4380 TGGGGCAGGGGACCGCAGAGCTCCGTGGGGAAGCTCAATCCCATCTTGCACAAGACAG 4439
Db 4425 TGGGGCAGGGGACCGCAGAGCTCCGTGGGGAAGCTCAATCCCATCTTGCACAAGACAG 4484
Qy 4440 CTTTGTCCAGCTGTCCACATTTGAGTCTGCTCCCGGGAGAGACCCCGGGCCCCCA 4499
Db 4485 CTTTGTCCAGCTGTCCACATTTGAGTCTGCTCCCGGGAGAGACCCCGGGCCCCCA 4544
Qy 4500 GCACATAAAGAACTGCGAGCCTTGGTACTGCGAGAGTCTGGGTGTAGAGAACTTTTGTAA 4559
Db 4545 GCACATAAAGAACTGCGAGCCTTGGTACTGCGAGAGTCTGGGTGTAGAGAACTTTTGTAA 4604
Qy 4560 GCAATAAAGTTTGGGTGATGACAAATGTT 4589
Db 4605 GCAATAAAGTTTGGGTGATGACAAATGTT 4634

RESULT 4

US-60-397-567
; Sequence 567, Application US/60659397
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERPERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/659,397
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 567
; LENGTH: 4458
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-659-397-567

Query Match 92.3%; Score 4243.2; DB 15; Length 4458;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 4343; Conservative 6; Mismatches 39; Indels 10; Gaps 7;

Qy 194 GTGTTTGGTCTCTCTCAGGATGAGCACAAAGCCTGGGAGATGCGAGTGGAAATGGCT 253
Db 69 GAGATCGGGCTTGGCCGGATGAGCACAAAGCCTGGGAGATGCGAGTGGAAATGGCT 128
Qy 254 GCCCAGGTGCCCCCTGGCTCAGCAGTGGGGCAGCAGAGAACTCCCCAAGCCAAAGGAGA 313
Db 129 GCCCAGGTGCCCCCTGGCTCAGCAGTGGGGCAGCAGAGAACTCCCCAAGCCAAAGGAGA 188
Qy 314 AGACCCCGCACTGGGGGAAGAAAACAGAGCTCGTCTACAAGTTAGGCCGTGGAGAAGA 373
Db 189 AGACCCCGCACTGGGGGAAGAAAACAGAGCTCGTCTACAAGTTAGGCCGTGGAGAAGA 248
Qy 374 GCCCTGTCTTCGCGAAAGTGGGAGATCCCTGAATGAGCTGATTAACCAAGGCCACAGCCA 433
Db 249 GCCCTGTCTTCGCGAAAGTGGGAGATCCCTGAATGAGCTGATTAACCAAGGCCACAGCCA 308
Qy 434 AGGAAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATCATGCCAGGCTGAGTGTAGA 493
Db 309 AGGAAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATCATGCCAGGCTGAGTGTAGA 368
Qy 494 ATAGCCAAAGGTTAGCCGCCACCTTTTCAGAAACGATTTTCATCGCTGGGTCAAACAGT 553

Db 369 ATAGCCAGAGTTTCAGCCCCACCTTTTCAGAACGCAATTTTCATCGCTGGGTCCAAACAGT 428
Qy 554 ACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAATGTGGCCCATCTACTACAGAGGCA 613
Db 429 ACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAATGTGGCCCATCTACTACAGAGGCA 488
Qy 614 AAATGGCCCGTGTGTGGAAAGGGAAGCTTCGACGCAAAAGCCCGGAGAAACGGGAAGA 673
Db 489 AAATGGCCCGTGTGTGGAAAGGGAAGCTTCGACGCAAAAGCCCGGAGAAACGGGAAGA 548
Qy 674 AGAAGAGCTCAAAAGTCCCTGGCTCATGACGAGTGGCTTCGCCCAAAACCCCTCCCAAGGA 733
Db 549 AGAAGAGCTCAAAAGTCCCTGGCTCATGACGAGTGGCTTCGCCCAAAACCCCTCCCAAGGA 608
Qy 734 CCCCTGACGAGAGAGTGCACCATCCACATGCCAGTGCAGGAGGATGATCTCCACTCGCGCGCC 793
Db 609 CCCCTGACGAGAGAGTGCACCATCCACATGCCAGTGCAGGAGGATGATCTCCACTCGCGCGCC 668
Qy 794 CATATGTTAGAAACACCCCGCAGTTCAACCAAGCTCTGMAAGGAACCAAGGCTTGGGCAC 853
Db 669 CATATGTTAGAAACACCCCGCAGTTCAACCAAGCTCTGMAAGGAACCAAGGCTTGGGCAC 728
Qy 854 TCTGTTTAAAGCAGTTGGCGAGGCGCTTACGGCCGGCTCTGCCCTCGATCAGAACTCCACA 913
Db 729 TCTGTTTAAAGCAGTTGGCGAGGCGCTTACGGCCGGCTCTGCCCTCGATCAGAACTCCACA 788
Qy 914 AACTGATCAGCCCTTGCATGTCTGAACACACGTGTGGAACCTGCACACCCCGCAGGACG 973
Db 789 AACTGATCAGCCCTTGCATGTCTGAACACACGTGTGGAACCTGCACACCCCGCAGGACG 848
Qy 974 GAGGCCCTCGCCCTGGCCAGCACACCTTCCCTATAGCAGACTGCTCATCCCTTCC 1033
Db 849 GAGGCCCTCGCCCTGGCCAGCACACCTTCCCTATAGCAGACTGCTCATCCCTTCC 908
Qy 1034 CATTCACCTCTCCAGCCTGGGAAACCTCAACCTCTGGAGTCTCTCTGGGCAAACTGG 1093
Db 909 CATTCACCTCTCCAGCCTGGGAAACCTCAACCTCTGGAGTCTCTCTGGGCAAACTGG 968
Qy 1094 CCTGTGTAGACAGCCAGAAACCTTTGGCTGACCCACACCTGAGCAAACTGGCCTGTGTAG 1153
Db 969 CCTGTGTAGACAGCCAGAAACCTTTGGCTGACCCACACCTGAGCAAACTGGCCTGTGTAG 1028
Qy 1154 ACAGTCCAAAGCCCTGCTGGCCCAACACCTGTGAGCCAGCCAGTGTCTGTGGTGCC 1213
Db 1029 ACAGTCCAAAGCCCTGCTGGCCCAACACCTGTGAGCCAGCCAGTGTCTGTGGTGCC 1088
Qy 1214 ATGAGAGATTTCTGTGGAGGAATACCTAGTGTGATGCTCTGCAAGGCAAGCTGAGCTCA 1273
Db 1089 ATGAGAGATTTCTGTGGAGGAATACCTAGTGTGATGCTCTGCAAGGCAAGCTGAGCTCA 1148
Qy 1274 GCCAGGCCCAAGCCTGACAGCCTGGCCAAAGACTTGGGCAAGCAGCGGGCTTCCAGATCCC 1333
Db 1149 GCCAGGCCCAAGCCTGACAGCCTGGCCAAAGACTTGGGCAAGCAGCGGGCTTCCAGATCCC 1208
Qy 1334 GGGAGCCAGCCCAAACTGAGGACAAACGAGGTGTCTGCTCACTGAGAACTCAAGC 1393
Db 1209 GGGAGCCAGCCCAAACTGAGGACAAACGAGGTGTCTGCTCACTGAGAACTCAAGC 1268
Qy 1394 CAGTGGATATGATGACGAGAAAGTCCACTGGGCCACGACACAGCTCGGCTGGGCA 1453
Db 1269 CAGTGGATATGATGACGAGAAAGTCCACTGGGCCACGACACAGCTCGGCTGGGCA 1328
Qy 1454 GAGGCTCTCTCGGAGAGGTGCACAGGATGGAGGAACAAGCAGTGGCTTCCAGTGGCTG 1513
Db 1329 GAGGCTCTCTCGGAGAGGTGCACAGGATGGAGGAACAAGCAGTGGCTTCCAGTGGCTG 1388
Qy 1514 TCAAAAAGGTGCGCTGGAGATTTTCGGGCAGAGGAGCTGATGGCATGTGCGAGATTGA 1573
Db 1389 TCAAAAAGGTGCGCTGGAGATTTTCGGGCAGAGGAGCTGATGGCATGTGCGAGATTGA 1448
Qy 1574 CCTCACCAGAAATGCTCCCTTTGTATGAGCTGTGAGAGAGGCGCTTGGGTCAACATCT 1633
Db 1449 CCTCACCAGAAATGCTCCCTTTGTATGAGCTGTGAGAGAGGCGCTTGGGTCAACATCT 1508

Qy 1634 TCATGGAGCTGCTGGAAAGGTGGTCTCCCTGGGCCAGCTGGTCAAGGAGCAGGGCTGTCTCC 1693
Db 1509 TCATGGAGCTGCTGGAAAGGTGGTCTCCCTGGGCCAGCTGGTCAAGGAGCAGGGCTGTCTCC 1568
Qy 1694 CAGAGGACCGGCCCTCTACTACTGGGCCAGGCCCTGGAGGCTCTGGAACTACCTCACT 1753
Db 1569 CAGAGGACCGGCCCTCTACTACTGGGCCAGGCCCTGGAGGCTCTGGAACTACCTCACT 1628
Qy 1754 CACGAAGATTTTCATGCTGGGACGTCAAAAGCTGACAACTGCTCTCTGCTCCAGCGATGGGA 1813
Db 1629 CACGAAGATTTTCATGCTGGGACGTCAAAAGCTGACAACTGCTCTCTGCTCCAGCGATGGGA 1688
Qy 1814 GCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGATGGCTCTGGAA 1873
Db 1689 GCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGATGGCTCTGGAA 1748
Qy 1874 AGTCTTGTCTCACAGGGGACTACATCCCTGGCACAGAGACCCACATGGCTCCGAGGTGG 1933
Db 1749 AGTCTTGTCTCACAGGGGACTACATCCCTGGCACAGAGACCCACATGGCTCCGAGGTGG 1808
Qy 1934 TGCTGGCAGGAGCTGCGACCGCCAAAGTGGATGTCTGGAGCAGCTGTGTATGATGCTGC 1993
Db 1809 TGCTGGCAGGAGCTGCGACCGCCAAAGTGGATGTCTGGAGCAGCTGTGTATGATGCTGC 1868
Qy 1994 ACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTCTGCCTCAAGA 2053
Db 1869 ACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTCTGCCTCAAGA 1928
Qy 2054 TTGCCAGGAGCTCCGCCCTGTGAGGAGATCCACCCCTCTGCGGCCCTCTCACAGCCC 2113
Db 1929 TTGCCAGGAGCTCCGCCCTGTGAGGAGATCCACCCCTCTGCGGCCCTCTCACAGCCC 1988
Qy 2114 AGGCCATCAAGAGGGCTGAGGAAAGAGCCATCCACCGGTCTGTGACGCGAGCTGG 2173
Db 1989 AGGCCATCAAGAGGGCTGAGGAAAGAGCCATCCACCGGTCTGTGACGCGAGCTGG 2048
Qy 2174 GAGGAAAGGTGAACCCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCTTTGGAGGGAG 2233
Db 2049 GAGGAAAGGTGAACCCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCTTTGGAGGGAG 2108
Qy 2234 AATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAAGACCCCTCCATG 2293
Db 2109 AATATAAAGAACCAAGAYATCCACCGCCAAATCAAGCCAAATACCAAGACCCCTCCATG 2168
Qy 2294 CCACGCGAGAGAGCTTTCCGCAAGGGCCCGAGGGCCCGGCGCCAGCTGAGGAGACAAAG 2353
Db 2169 CCACGCGAGAGAGCTTTCCGCAAGGGCCCGAGGGCCCGGCGCCAGCTGAGGAGACAAAG 2228
Qy 2354 GCAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAGAGCCCAACAAAGTCTC 2413
Db 2229 GCAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAGAGCCCAACAAAGTCTC 2288
Qy 2414 CTCCCTTGACTTTGAGCAAGGAGAGTCTGGATGTGGGAAACCTTACCTGTCTCTCC 2473
Db 2289 CTCCCTTGACTTTGAGCAAGGAGAGTCTGGATGTGGGAAACCTTACCTGTCTCTCC 2348
Qy 2474 TGGAGCCAGCCCTGCGCAGAAACCCAGAGCTCACAGAGCGGAAAGCAACCGTCCCGGAGC 2533
Db 2349 TGGAGCCAGCCCTGCGCAGAAACCCAGAGCTCACAGAGCGGAAAGCAACCGTCCCGGAGC 2408
Qy 2534 AGGAATCGCAGCAGCTGGAAATAGAAATTTCTCTCAACAGCCTCTCCAGCCATTTTCTC 2593
Db 2409 AGGAATCGCAGCAGCTGGAAATAGAAATTTCTCTCAACAGCCTCTCCAGCCATTTTCTC 2468
Qy 2594 TGGAGGAGCAGGACAAATTTCTCTGCTCTGAGATCGACAGCTCTCTCTGTGCGATG 2653
Db 2469 TGGAGGAGCAGGACAAATTTCTCTGCTCTGAGATCGACAGCTCTCTCTGTGCGATG 2528
Qy 2654 ACAGTGAAGAAACCCATCAAGGCGCTCTCAAGCTCGCGGACACCTTGAGCTCAGGCG 2713
Db 2529 ACAGTGAAGAAACCCATCAAGGCGCTCTCAAGCTCGCGGACACCTTGAGCTCAGGCG 2588


```
; PRIOR FILING DATE: 1996-04-02
; PRIOR APPLICATION NUMBER: IL 119133
; PRIOR FILING DATE: 1996-08-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 2631
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1081)..(1081)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1102)..(1102)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1120)..(1120)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1125)..(1125)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1129)..(1129)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1135)..(1135)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1146)..(1146)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1170)..(1170)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1180)..(1180)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1188)..(1188)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1208)..(1208)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1239)..(1239)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1248)..(1249)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1278)..(1278)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1297)..(1297)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1310)..(1310)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1322)..(1322)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1345)..(1345)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1409)..(1409)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1423)..(1423)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1445)..(1445)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1452)..(1452)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1459)..(1459)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1478)..(1478)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1498)..(1498)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1507)..(1508)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1520)..(1520)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1534)..(1534)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1540)..(1540)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1546)..(1546)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1557)..(1557)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1713)..(1713)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1895)..(1895)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1900)..(1900)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
```



```
; NAME/KEY: misc_feature
; LOCATION: (1934)..(1934)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1942)..(1942)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1951)..(1951)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1962)..(1962)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1967)..(1967)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1974)..(1974)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1984)..(1984)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1988)..(1988)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1994)..(1994)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2005)..(2005)
; OTHER INFORMATION: n is a, c, g, or t

Query Match 40.3%; Score 1851.2; DB 6; Length 2631;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2335; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

QY 2099 CCCCTCTCACAGCCAGCCGCTCCAGAGGGGCTCAGGAAGAGCCCATCCACCGCGTGT 2158
DB 1 CCCCCTCTCACAGCCAGCCGCTCCAGAGGGGCTCAGGAAGAGCCCATCCACCGCGTGT 60
QY 2159 CTGCAGCGAGCTGGGAGGAGGTCAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 2218
DB 61 CTGCAGCGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120
QY 2219 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCAATCAAGCCAAATTACC 2278
DB 121 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCAATCAAGCCAAATTACC 180
QY 2279 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTCCCAAGGGCCCGAGGGCCCGGCCAG 2338
DB 181 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTCCCAAGGGCCCGAGGGCCCGGCCAG 240
QY 2339 CTGAGGAGACAACAGCGAGAGCCCTTAAGCTCCAGCTCCTCTCCACCAAGAGCCCGCCAG 2398
DB 241 CTGAGGAGACAACAGCGAGAGCCCTTAAGCTCAGGCTCCTCTCCACCAAGAGCCCGCCAG 300
QY 2399 AGCCAAACAAGTCTCCTCTCTTGAATTTGAGCAAGAGAGGTCTGGGATGTGGGAACCTT 2458
DB 301 AGCCAAACAAGTCTCCTCTCTTGAATTTGAGCAAGAGAGGTCTGGGATGTGGGAACCTT 360
QY 2459 TACCTCTGTCTCCCTGGAGCCAGCCCTGCGCAGAAACCCAGAGTCAACAGAGCGGAAG 2518
DB 361 TACCTCTGTCTCCCTGGAGCCAGCCCTGCGCAGAAACCCAGAGTCAACAGAGCGGAAG 420
QY 2519 CAACCGTCCCGAGCAGGAATGCAGCAGCTGGAAATAGAAATTATTCCTCAACAGCCTGT 2578
```

Db 421 CAACCGTCCCGAGCAGGAATGCAGCAGCTGGAAATAGAAATTATTCCTCAACAGCCTGT 480

QY 2579 CCAGGCCATTTCTCTGGAGCAGGAGCAAAATCTCTCGTCTGCTCAGCATCGACAGCC 2638

Db 481 CCCAGCCATTTCTCTGGAGCAGGAGCAAAATCTCTCGTCTGCTCAGCATCGACAGCC 540

QY 2639 TCTCCCTGTCGATGACAGTGAAGAACCCATCAAAAGCCCTCTCAAAAGCTCGCGGGACA 2698

Db 541 TCTCCCTGTCGATGACAGTGAAGAACCCATCAAAAGCCCTCTCAAAAGCTCGCGGGACA 600

QY 2699 CCCTGAGCTCAGGCGTACATCTCTGGAGCAGCAGCGGCTCGAAGCTCCAGCTGA 2758

Db 601 CCCTGAGCTCAGGCGTACATCTCTGGAGCAGCAGCGGCTCGAAGCTCCAGCTGA 660

QY 2759 ACATGGTCTGGCCGGGGGGCCACCGACACCCCAAGCTATTTCATGGTGAAG 2818

Db 661 ACATGGTCTGGCCGGGGGGGGCCACCGACACCCCAAGCTATTTCATGGTGAAG 720

QY 2819 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCGGAGTTCACACGGGTCAAAG 2878

Db 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCGGAGTTCACACGGGTCAAAG 780

QY 2879 TGGGAGACATCGCCATCGCATCAGACGCCAGATCCAGCTGCAGCTTCAGCTTGGTCA 2938

Db 781 TGGGAGACATCGCCATCGCATCAGACGCCAGATCCAGCTGCAGCTTCAGCTTGGTCA 840

QY 2939 CCAAGACGGCAGCCTTTCGCTACGACATGGAGTGCAGACTCGGGCATCGACCTGC 2998

Db 841 CCAAGACGGCAGCCTTTCGCTACGACATGGAGTGCAGACTCGGGCATCGACCTGC 900

QY 2999 AGTGACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGC 3058

Db 901 AGTGACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGC 960

QY 3059 TGGAGAACAGGCCCTTAACCTTCCACCGCGGTCCACACTGCGGAAAGCAGCT 3118

Db 961 TGGAGAACAGGCCCTTAACCTTCCACCGCGGTCCACACTGCGGAAAGCAGCT 1020

QY 3119 TCCTGCTCGGTGACGATGCTCCCTGAAACACAGGCTCAGCGTTCACAGGGATTC 3176

Db 1021 TCCTGCTCGGTGACGATGCTCCCTGAAACACAGGCTCAGCGTTCACAGGGATTC 1080

QY 3177 GCCAGCCCCCGGTCA--CAGTGGGAACAGGGGCTCG-----CAGCAGCAAGGTGGGG 3229

Db 1081 NCCAGCCCCCGGTCA--CAGTGGGAACAGGGGCTCG-----CAGCAGCAAGGTGGGG 1140

QY 3230 GCAAGCAAAATGCTCTCCAGGATTCACCTGAGCCCTGCCCCAC-----CCTGTGA 3283

Db 1141 CAAGCAAAATGCTCTCCAGGATTCACANCTGAGCCCTGCCCANCCCTGCTGAADA 1200

QY 3284 AAAACATCCGCCAGTGAAGACAGAGGAGGTGGC---AGGAGTTACTCGGGAA 3339

Db 1201 AAACATNCCGCAAGTGAAGACAGAGGAGGTGGC---AGGAGTTACTCGGGAA 1260

QY 3340 ACAGAACAGGATCTTT--TTCTGCCCTGCTCAGT-CGAGTTGGCTGA--CCGCTTG 3395

Db 1261 ACAGAACAGGATCTTT--TTCTGCCCTGCTCAGT-CGAGTTGGCTGA--CCGCTTG 1320

QY 3396 GATCAGTGAACATTTGTTGGCAGA-CAGGGAGAGCAGCTTCAGCTGGGTGAGAGGG 3454

Db 1321 ANTCAGTGAACATTTGTTGGCAGANAGGGAGAGCAGCTTCAGCTGGGTGAGAGGG 1380

QY 3455 GTGGCGAGCCCTTCGGCCCTCAGCT--CCAGGCTGCTGTG-AGAGTGTCAAGTGTGA 3512

Db 1381 GTGGCGAGCCCTTCGGCCCTCAGCT--CCAGGCTGCTGTG-AGAGTGTCAAGTGTGA 1440

QY 3513 AGGGCCCAAACTC---AGGTTCAAGTGAAGAACAGGT-CAGCAGGTATGCCCGCCGTAG 3568

Db 1441 AGGGCCCAAACTC---AGGTTCAAGTGAAGAACAGGT-CAGCAGGTATGCCCGCCGTAG 1500

QY 3569 GTTAA-----GGGGGCGCTCTAAACCCCTTGGCTGGCTCA---CCTGGCAGCTCA-CCC 3620

Db 58956 GTGTAGGGGAAAGAAATGCTGACCCCTGGAAAGGCTCCCTGGTAGAATACACCACTTT 59015
QY 3688 TCAGGTGTGTGCAACACAGAGTCTTGAGTTGACCTCTGGTTTACGCAAGGACCAAGAGAGG 3747
Db 59016 TCAGGTGTGTGCAACACAGAGTCTTGAGTTGACCTCTGGTTTACGCAAGGACCAAGAGAGG 59075
QY 3748 TGTGTAAGTGAAGTGTGTTCTCAGTCCCAAGACATGTCGCCCTTTGCTGTGCTACCACT 3807
Db 59076 TGTGTAAGTGAAGTGTGTTCTCAGTCCCAAGACATGTCGCCCTTTGCTGTGCTACCACT 59135
QY 3808 CTTCCCCAGAGCAGAGGCCCGAGGCCCTTCAGGCCCAAGCACTGCCCAAGCACTGCTGG 3867
Db 59136 CTTCCCCAGAGCAGAGGCCCGAGGCCCTTCAGGCCCAAGCACTGCCCAAGCACTGCTGG 59195
QY 3868 CACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTGAAGGATATGCAAGATATGCTGACAGGAACAG 3927
Db 59196 CACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTGAAGGATATGCAAGATATGCTGACAGGAACAG 59255
QY 3928 TCTGTGGATGGAATGATCAGTGTCTAAGGAAGACAGAGAGAGAGC-TCCGGCGCCCC 3986
Db 59256 TCTGTGGATGGAATGATCAGTGTCTAAGGAAGACAGAGAGAGAGC-TCCGGCGCCCC 59315
QY 3987 AGCCCACTATCAGTGTCCAGGCTGCTGGTTCCCAAGACAGCACTCAGCATCACATGA 4046
Db 59316 AGCCCACTATCAGTGTCCAGGCTGCTGGTTCCCAAGACAGCACTCAGCATCACATGA 59375
QY 4047 CACTCACTGCTGCCCTGCCCTGCGCAGAGGGTACTGCCAGCGCACTTTGCACTCTGATG 4106
Db 59376 CACTCACTGCTGCCCTGCCCTGCGCAGAGGGTACTGCCAGCGCACTTTGCACTCTGATG 59435
QY 4107 ACCTCAAGACATTTTCATGGTGCCTCTGGCAGGGCAGGGCAGGGCAGTGCATGTAG 4166
Db 59436 ACCTCAAGACATTTTCATGGTGCCTCTGGCAGGGCAGGGCAGGGCAGTGCATGTAG 59495
QY 4167 GAGCATAGCAAGCCAGGATGGGTGAAGGACACAGTCTTGACGTCTCCACATGCATG 4226
Db 59496 GAGCATAGCAAGCCAGGATGGGTGAAGGACACAGTCTTGACGTCTCCACATGCATG 59555
QY 4227 TGACTCCTCAAACTCTTCCAGATTTCTTAAGATAGCACCCCTTCCCAATTCGCCCA 4286
Db 59556 TGACTCCTCAAACTCTTCCAGATTTCTTAAGATAGCACCCCTTCCCAATTCGCCCA 59615
QY 4287 GCTTAGCCTTTCTCCAGGGAGGCTACTCAGGAATCACTAGATTAATAACAGTGTGA 4346
Db 59616 GCTTAGCCTTTCTCCAGGGAGGCTACTCAGGAATCACTAGATTAATAACAGTGTGA 59675
QY 4347 ATCGTCAGGGGTGCTGTAGCCTCAACCTCTGGGCGAGGGGCGGAGACTCCGTG 4406
Db 59676 ATCGTCAGGGGTGCTGTAGCCTCAACCTCTGGGCGAGGGGCGGAGACTCCRTG 59735
QY 4407 GGAGAAGCTCATTTCCACATCTTGCAAGACAGACGCTTTGTCCAGCTGTCCACATTGAGTC 4466
Db 59736 GGAGAAGCTCATTTCCACATCTTGCAAGACAGACGCTTTGTCCAGCTGTCCACATTGAGTC 59795
QY 4467 AGACTGCTCCGGGAGAGAGCCCCCGCCCCAGCACATAAAGAACTCCAGACCTTGGTAC 4526
Db 59796 AGACTGCTCCGGGAGAGAGCCCCCGCCCCAGCACATAAAGAACTCCAGACCTTGGTAC 59855
QY 4527 TGCAGAGTCTGGGTGTAGAGCACTTTGTAAAGCAATAAAGTTTGGGGTGTAGCAAT 4586
Db 59856 TGCAGAGTCTGGGTGTAGAGCACTTTGTAAAGCAATAAAGTTTGGGGTGTAGCAAT 59915
QY 4587 GTT 4589
Db 59916 GTT 59918

RESULT 7

US-10-450-763-15295/c
; Sequence 15295, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 15295
; LENGTH: 2760
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1282)..(1566)
; OTHER INFORMATION: 73% homologous to Mus musculus Pro-Pol-dUTPase
; OTHER INFORMATION: polypeptide, accession number Y12713, Smith-Waterman Score=381.
US-10-450-763-15295

Query Match 32.1%; Score 1474.4; DB 8; Length 2760;
Best Local Similarity 73.4%; Pred. No. 0;
Matches 2370; Conservative 0; Mismatches 6; Indels 852; Gaps 4;

QY 232 ATGGCAGTGTGGAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGCGAGCAGAG 291
Db 2760 ATGGCAGTGTGGAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGCGAGCAGAG 2701
QY 292 GAACTCCCAAGCCAAAGGAGAGACGCGCCCACTGGGGAAGAAACACAGCTCCGTCTTAC 351
Db 2700 GAACTCCCAAGCCAAAGGAGAGACGCGCCCACTGGGGAAGAAACACAGCTCCGTCTTAC 2641
QY 352 AAGCTTAGGCGCTGGAGAGAGCCCTGTGTGTCGGAAAGTGGAGATCCTGAATGAC 411
Db 2640 AAGCTTAGGCGCTGGAGAGAGCCCTGTGTGTCGGAAAGTGGAGATCCTGAATGAC 2581
QY 412 GTGATTACCAAGGCACAGCAAGGCTCCGAGGCGAGGCGCAGCTCCTATC 471
Db 2580 GTGATTACCAAGGCACAGCAAGGCTCCGAGGCGAGGCGCAGCTCCTATC 2521
QY 472 ATCGCCCAAGGCTGAGT----- 487
Db 2520 ATCGCCCAAGGCTGAGTGTAGGGGCTAGGGTGTGTTGGGCACGCGGCGAGGCGCTTACC 2461
QY 488 ----- 487
Db 2460 ATCTTCCACCTAGAGGTGGCCCCA CCCCCAGGAATGAGCACCGTCTCTTGAGCCTGCTG 2401
QY 488 -----GTGAGAAATAGCAAGAGTTCAGGCCCAACC 516
Db 2400 GGCAGTCTGGGGGGGAGGCTGCTGAGGGTGAAGATAGCAAGAGTTCAGGCCCAACC 2341
QY 517 TTTTCAGAACGCAATTTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGT 576
Db 2340 TTTTCAGAACGCAATTTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGT 2281
QY 577 CAGATCCCAACAAATGTGGCCCATGCTACAGAGGGCAAAATGGCCCGTGTGTGTGGAG 636
Db 2280 CAGATCCCAACAAATGTGGCCCATGCTACAGAGGGCAAAATGGCCCGTGTGTGTGGAG 2221
QY 637 GGAAGGCTGCGACAAAGCCCGGAAGAAACGGAAGAGAGAGTCAAAAGTCCCTGGCT 696
Db 2220 GGAAGGCTGCGACAAAGCCCGGAAGAAACGGAAGAGAGAGTCAAAAGTCCCTGGCT 2161
QY 697 CATCAGAGTGGCTTTGGCCAAACCCCTCCCAAGGACCCCTGAGCAGGAGAGTGTGACC 756
Db 2160 CATCAGAGTGGCTTTGGCCAAACCCCTCCCAAGGACCCCTGAGCAGGAGAGTGTGACC 2101
QY 757 ATCCCAAGTGCAGGAGATGAGTCTCCACTCGGCCGCCCATATGTTAGNAACACCCCCGAG 816
Db 2100 ATCCCAAGT----- 2092

Qy	817	TTCAACAAGCCTCTGAAGGAACCAAGGCCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAG	876
Db	2091	-----	2092
Qy	877	GGCCTACGGCCGGCTCTGCCTCGATCAGAACTCCACAACTGATCATGCCCTTGCATGT	936
Db	2091	-----	2092
Qy	937	CTGAACCAACGTGTGGAAACTGCACCAACCCCAAGGACGGAGGCCCTGCCCCCTGCCCAACG	996
Db	2091	-----	2092
Qy	997	CACCCCTTCCCCTATAGCAGACTGCCCTCATGCCCTTCCCACTCCACCCCTCTCCAGCCCTGG	1056
Db	2091	-----	2083
Qy	1057	AAACCTCACCTCTGGAGTCTCTTCTGGGCATAACTGGCCTGTGTAGACAGCAGAAACCC	1116
Db	2082	AAACCTCACCTCTGGAGTCTCTTCTGGGCATAACTGGCCTGTGTAGACAGCAGAAACCC	2023
Qy	1117	TTGCCTGACCCACACTTGAGCAAACTGGCCTGTGTAGACAGTCAAAAGCCCTGCCTGGC	1176
Db	2022	TTGCCTGACCCACACTTGAGCAAACTGGCCTGTGTAGACAGTCCAAAGCCCTGCCTGGC	1963
Qy	1177	CCACACCTGGAGCCACGCTGCCTGTCTGTGGTGTGCCCATGAGAAGTTTTCTGTGGAGAA	1236
Db	1962	CCACACCTGGAGCCACGCTGCCTGTCTGTGGTGTGCCCATGAGAAGTTTTCTGTGGAGAA	1903
Qy	1237	TACCTAGTGCATGCTCTGCAAGGCAGCGTGAGCTCAAGCCAGGCCACAGCCTGACCAGC	1296
Db	1902	TACCTAGTGCATGCTCTGCAAGGCAGCGTGAGCTCAGCCAGGCCACAGCCTGACCAGC	1843
Qy	1297	CTGGCCAAAGACCTGGGCAGCACGCGGCTCCAGATCCCGGAGGCCAGCCCCAAAACCTGAG	1356
Db	1842	CTGGCCAAAGACCTGGGCAGCAAGGGCTCCAGATCCCGGAGGCCAGCCCCAAAACCTGAG	1783
Qy	1357	GACAAACAGGGTGTCTCTCACTGAGAAACTCAAGCCAGTGGATTATGATACCCGAGAA	1416
Db	1782	GACAAACAGGGTGTCTCTCACTGAGAAACTCAAGCCAGTGGATTATGATACCCGAGAA	1723
Qy	1417	GAAGTCCACTGGGCCACGACCAAGCTCCGCTCGGGCAGAGGCTCCTTCGGAGAGGTGCAC	1476
Db	1722	GAAGTCCACTGGGCCACGACCAAGCTCCGCTCGGGCAGAGGCTCCTTCGGAGAGGTGCAC	1663
Qy	1477	AGGATGAGGACAAAGCAGACTGGGTTTCAAGTGGCTGTCAAAA-----	1519
Db	1662	AGGATGAGGACAAAGCAGACTGGGTTTCAAGTGGCTGTCAAAAAGGACAGTCGCAAGGGTG	1603
Qy	1520	-----	1519
Db	1602	GCGGTGGGTCTGACTTGTGAACAGAACTTGGCTATCATAGCCACCTGAGGTCGGCTGGGA	1543
Qy	1520	-----	1519
Db	1542	ACTGGAAGGAGCAGCCTGTACCGTGTCTTGGCCTGGCTGCTGCCCTTGGCATGTGTGTA	1483
Qy	1520	-----	1519
Db	1482	TGTGAGAGGCAGTGTGAGAGSACACAGTGGCCCTCGGTGTCTCAATACAGCCAGGGCCTG	1423
Qy	1520	-----	1527
Db	1422	AGGGCAAGTCCCAACGTGGAGAGTGGTTGGAGCGTCAAGAAATTCACACTGAGTGGCGG	1363
Qy	1528	CTGGAAGTATTTTCGGGCAGAGGAGCTGATGSCATGTGCAGGATTTGACCTCACCCAGAATT	1587
Db	1362	CTGGAAGTATTTTCGGGCAGAGGAGCTGATGSCATGTGCAGGATTTGACCTCACCCAGAATT	1303
Qy	1588	GTCCCTTTGTATGAGACTGTGAGAGAAGGGCCTTGGGTCAACATCTTCATATGGAGCTGCTG	1647
Db	1302	GTCCCTTTGTATGAGACTGTGAGAGAAGGGCCTTGGGTCAACATCTTCATATGGAGCTGCTG	1243

QY	1648	GAAGGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAGGGCTGTCTCCAGAGGACCGGGGCC	1707
DB	1242	GAAGGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAGGGCTGTCTCCAGAGGACCGGGGCC	1183
QY	1708	CTGTACTACTCTGGGCAGGCCCTGGAAGGTCTCGAATACTCCACATCAGAGGATTTCTG	1767
DB	1182	CTGTACTACTCTGGGCAGGCCCTGGAAGGTCTCGAATACTCCACATCAGAGGATTTCTG	1123
QY	1768	CATGGGACGCTCAAAAGCTGACAACAGTGTCTCTGTCCAGCGATGGGACCAACGAGCCCTC	1827
DB	1122	CATGGGACGCTCAAAAGCTGACAACAGTGTCTCTGTCCAGCGATGGGACCAACGAGCCCTC	1063
QY	1828	TGTGACTTTGGCCATGCTGTGTCTTTCAACCTGATGCGCTGGGAAAGTCTCTTGCTCACA	1887
DB	1062	TGTGACTTTGGCCATGCTGTGTCTTTCAACCTGATGCGCTGGGAAAGTCTCTTGCTCACA	1003
QY	1888	GGGACCTACATCCCTGTGSCACAGAGACCCACATGGCTCCGAGGTGTGTGGGCAGGAGC	1947
DB	1002	GGGACCTACATCCCTGTGSCACAGAGACCCACATGGCTCCGAGGTGTGTGGGCAGGAGC	943
QY	1948	TGCGACGCCAAGGTGGATGTCTGGAGCAGCTGCTGTATGATGCTGCACATGCTCAACGGC	2007
DB	942	TGCGACGCCAAGGTGGATGTCTGGAGCAGCTGCTGTATGATGCTGCACATGCTCAACGGC	883
QY	2008	TGCCACCCCTGGACTCAGTTGTTCTTCGAGGGCCGCTCTGCCCTCAAGATTGCGCAGGACCT	2067
DB	882	TGCCACCCCTGGACTCAGTTGTTCTTCGAGGGCCGCTCTGCCCTCAAGATTGCGCAGGACCT	823
QY	2068	CCGCTGTGAGGGAGATCCACCCCTCTCGCGCCCTCTCACAGCCAGGCCATCCAAAG	2127
DB	822	CCGCTGTGAGGGAGATCCACCCCTCTCGCGCCCTCTCACAGCCAGGCCATCCAAAG	763
QY	2128	GGGCTGAGGAAAGAGCCCATCCACCGCGTGTCTGCAGCGGAGCTGGGAGGGAAGTGAAC	2187
DB	762	GGGCTGAGGAAAGAGCCCATCCACCGCGTGTCTGCAGCGGAGCTGGGAGGGAAGTGAAC	703
QY	2188	CGGGCACTACAGCAAGTGGGAGTCTGAAGAGCCCTTTGGAGGGGAGAAATAAAGAACCA	2247
DB	702	CGGGCACTACAGCA-----	689
QY	2248	AGACATCCACGGCCAAATCAAGCCATTACCAACGACCTCTCATGCCAGCCGAGAGAG	2307
DB	688	-----	689
QY	2308	CTTTTCGCAAGGGCCCCAGGGCCCCCGGCAGCTGAGGAGACAAAGCAGAGCCCCCTAAG	2367
DB	688	-----	689
QY	2368	CTCAGGCTCTCTTCCACAGAGCCCCCAGAGCCAAACAAAGTCTCTCTCCCTTGACTTTG	2427
DB	688	-----AGAGCCCCCAGAGCCAAACAAAGTCTCTCTCCCTTGACTTTG	649
QY	2428	AGCAAGAGGAGTCTGGATGTGGGAACCTTTACCTCTGTCTCTCTCCCTGGAGCCAGCCCT	2487
DB	648	AGCAAGAGGAGTCTGGGATGTGGGAACCTTTACCTCTGTCTCTCTCCCTGGAGCCAGCCCT	589
QY	2488	GCCAGAAAACCCAGCTCACACAGAGCGGAAAGCAACCGTCCCGAGCAGGAATCTGCAGCAG	2547
DB	588	GCCAGAAAACCCAGCTCACACAGAGCGGAAAGCAACCGTCCCGAGCAGGAATCTGCAGCAG	529
QY	2548	CTGGAATAGAAATTTCTTCTCAACAGCCTGTCCAGCCATTTTTCTCTGGAGGAGCAGGAG	2607
DB	528	CTGGAATAGAAATTTCTTCTCAACAGCCTGTCCAGCCATTTTTCTCTGGAGGAGCAGGAG	469
QY	2608	CAAAATTTCTGTGCTCTCAGCATGACAGCCCTCTCCCTGTCCGATGACAGTGAAGAAC	2667
DB	468	CAAAATTTCTGTGCTCTCAGCATGACAGCCCTCTCCCTGTCCGATGACAGTGAAGAAC	409
QY	2668	CCATCAAGGCCTCTCAAAAGCTCGCGGACACCTTGAGCTCAGCGGTACTCTCTGGAGC	2727
DB	408	CCATCAAGGCCTCTCAAAAGCTCGCGGACACCTTGAGCTCAGCGGTACTCTCTGGAGC	349
QY	2728	AGCCAGGCCAGGCTCGAAGCTCCAGCTGGAACATGCTGTGGCCCGGGGGCGGCCACCC	2787

Db 348 AGCCAGGCGGAGCTCGAAGCTCCAGCTGGAACATGCTGGCCCGGGGGCGGCCACC 289
QY 2788 GACACCCCAAGCTATTTCATAGGTGGAAGTCCAAATACAGTCTCTTAATGGTGAACAC 2847
Db 289 GACACCCCAAGCTATTTCATAGGTGGAAGTCCAAATACAGTCTCTTAATGGTGAACAC 229
QY 2848 CTGCACATCCGGGAGTTCCACCGGGTCAAAGTGGGAGACATCGCCACTGGCATCAGCAGC 2907
Db 228 CTGCACATCCGGGAGTTCCACCGGGTCAAAGTGGGAGACATCGCCACTGGCATCAGCAGC 169
QY 2908 CAGATCCCAAGCTGCGACCTTCAGCTTGGTACCAAAGACGGGCGGCGCTGTTGCTACGAC 2967
Db 168 CAGATCCCAAGCTGCGACCTTCAGCTTGGTACCAAAGACGGGCGGCGCTGTTGCTACGAC 109
QY 2968 ATGGAGGTGCCAGATCGGGGATCGACCTGAGTGCACACTGGCCCTCGATGGGAGCTTC 3027
Db 108 ATGGAGGTGCCAGATCGGGGATCGACCTGAGTGCACACTGGCCCTCGATGGGAGCTTC 49
QY 3028 GCCTGGAGCTGGAGGTCAAGCATGGCCAGCTGGGAGAACAGGCCCTAA 3075
Db 48 GCCTGGAGCTGGAGGTCAAGCATGGCCAGCTGGGAGAACAGGCCCTAA 1

RESULT 8

US-11-060-756-560
; Sequence 560, Application US/11060756
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 560
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-560

Query Match 29.3%; Score 1345.6; DB 13; Length 1400;
Best Local Similarity 98.6%; Pred. No. 5.3e-293;
Matches 1384; Conservative 5; Mismatches 9; Indels 6; Gaps 3;
QY 3195 AGTGGGAACCAAGGCTCGCAGCAGCAAGGTGGGGCAAGCAAGATGCTCCAGGATTT 3254
Db 1 AGTGGGAACCAAGGCTCGCAGCAGCAAGGTGGGGCAAGCAAGATGCTCCAGGATTT 60
QY 3255 CACACCTCAGCCCTGCCACCCCTGCTGAAATAACA-TCCGCCAGCTGAAGAGACAGAAG 3313
Db 61 CACACCTCAGCCCTGCCACCCCTGCTGAAATAACA-TCCGCCAGCTGAAGAGACAGAAG 120
QY 3314 GAGGATGGCAGGAGTTACTCTGGGAAACAAACAGGGATCTTTTCTGCCCTGCTCCAG 3373
Db 121 GAGGATGGCAGGAGTTACTCTGGGAAACAAACAGG---ATCTTCTGCCCTGCTCCAG 176
QY 3374 TCGAGTTGGCTGACCCGCTTGATTCAGTGACCAATTTGTGGCAGACAGGGGAGAGCAGC 3433
Db 177 TCGAGTTGGCTGACCCGCTTGATTCAGTGACCAATTTGTGGCAGACAGGGGAGAGCAGC 236
QY 3434 TTCCAGCTGGGTGAGAGGGGTGGCGAGCCCTTCGCGCCCTCACCCCTCAGGCTGCTG 3493
Db 237 TTCCAGCTGGGTGAGAGGGGTGGCGAGCCCTTCGCGCCCTCACCCCTCAGGCTGCTG 296
QY 3494 TGAGAGTGTCAAGTGTGAAGGGGCCAAACTCAGGTTTCAGTGCAGAACCCAGGTGAGCAGG 3553
Db 297 TGAGAGTGTCAAGTGTGAAGGGGCCAAACTCAGGTTTCAGTGCAGAACCCAGGTGAGCAGG 356
QY 3554 TATGCCGCCCGTAGGTTAAGGGGCCCTCTAAACCCCTTGGCTGGCCCTACCTGGGCCAG 3613

Db 357 TATGCCGCCCGTAGGTTAAGGGGCCCTCTAAACCCCTTGGCTGGCCCTCACCTGGCCAG 416
QY 3614 CTCACCCCTTTTGGTGTAGGGGAAAAAATGCTGACCTGAGAGAGGCTCCCTGGTAGA 3673
Db 417 CTCACCCCTTTTGGTGTAGGGGAAAAAATGCTGACCTGAGAGAGGCTCCCTGGTAGA 476
QY 3674 ATACACACACTTTTCAGGTTGTGCAACACAGAGTCTGAGTTGACCTCTGTTTTCAGCCA 3733
Db 477 ATACACACACTTTTCAGGTTGTGCAACACAGAGTCTGAGTTGACCTCTGTTTTCAGCCA 536
QY 3734 AGGACCAAGAGGTGTGTAAGTGAAGTGTCTCAGTCCCAGACATGTGCCCCCTTTGC 3793
Db 537 AGGACCAAGAGGTGTGTAAGTGAAGTGTCTCAGTCCCAGACATGTGCCCCCTTTGC 596
QY 3794 TGCTGGCTACACACTTTCCCGCAGAGCAGAGCCCGGAGCCCTTCAGGCCACGACCTGC 3853
Db 597 TGCTGGCTACACACTTTCCCGCAGAGCAGAGCCCGGAGCCCTTCAGGCCACGACCTGC 656
QY 3854 CCCAGACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGCAGGATATG 3913
Db 657 CCCAGACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGCAGGATATG 716
QY 3914 CCTGACAGGAACAGTCTGTGGATGACATGATCAGTGTCTAAAGAAAGCAGCAGAGAGA 3973
Db 717 CCTGACAGGAACAGTCTGTGGATGACATGATCAGTGTCTAAAGAAAGCAGCAGAGAGA 776
QY 3974 CG-TCCGGGCGCCAGCCCGCAGCTCAGTGTCCAGCGTGTCTGCCAGAGCAGCT 4032
Db 777 CGCTCCGGCGCCCGCAGCCCGCAGCTCAGTGTCCAGCGTGTCTGCCAGAGCAGCT 836
QY 4033 CAGCATCACACTGACACTCACCTCGCTGCCCTGGCGCAGAGGTACTGCCAGCGCAC 4092
Db 837 CAGCATCACACTGACACTCACCTCGCTGCCCTGGCGCAGAGGTACTGCCAGCGCAC 896
QY 4093 TTGCACTCTGATGACCTCAAAGCAGCTTTATGCTGCTGCTGCGAGGGCAGGCGAGG 4152
Db 897 TTGCACTCTGATGACCTCAAAGCAGCTTTATGCTGCTGCTGCGAGGGCAGGCGAGG 956
QY 4153 CAGTGACACTCTAGGAGATAGCAGCCAGAGATGGGGTGAAGGGAACACAGTCTTGAGC 4212
Db 957 CAGTGACACTCTAGGAGATAGCAGCCAGAGATGGGGTGAAGGGAACACAGTCTTGAGC 1016
QY 4213 TGTCCACATGCTGACTCTCTCAAACTCTTCAGAGATTTCTTAAGAAATAGCACCCCT 4272
Db 1017 TGTCCACATGCTGACTCTCTCAAACTCTTCAGAGATTTCTTAAGAAATAGCACCCCT 1076
QY 4273 TCCCAATGCCCCAGCTTAGCTCTTCTCCAGGGGAGCTACTCAGGACTCAGCTAGCAT 4332
Db 1077 TCCCAATGCCCCAGCTTAGCTCTTCTCCAGGGGAGCTACTCAGGACTCAGCTAGCAT 1136
QY 4333 TAAATCAGCTGTGAATCGTCAAGGGGTGTCTGCTAGCTCAACCTCTGGGGCAGGGGAC 4392
Db 1137 TAAATCAGCTGTGAATCGTCAAGGGGTGTCTGCTAGCTCAACCTCTGGGGCAGGGGAC 1196
QY 4393 GCCGAGACTCCGCTGGGAGAGCTCATTTCCCACTCTTGGCCAAAGCAGCTTTTTCAGCT 4452
Db 1197 GCCGAGACTCCGCTGGGAGAGCTCATTTCCCACTCTTGGCCAAAGCAGCTTTTTCAGCT 1256
QY 4453 GTCCACATTTAGTCAAGTGTCTCCCGGGGAGAGAGCCCGGCCCGCCAGCATAAAGAAC 4512
Db 1257 GTCCACATTTAGTCAAGTGTCTCCCGGGGAGAGAGCCCGGCCCGCCAGCATAAAGAAC 1316
QY 4513 TGCAGCTTGGTACTGAGAGTCTGGGTGTGAGAGAACTCTTTTGAAGCAATAAGTTTG 4572
Db 1317 TGCAGCTTGGTACTGAGAGTCTGGGTGTGAGAGAACTCTTTTGAAGCAATAAGTTTG 1376
QY 4573 GGGTGTGACAAATGTTAAAAAAA 4596
Db 1377 GGGTGTGACAAATGTTAAAAAAA 1400

; Sequence 4832, Application US/11060756
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TARGET GENES:
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4832
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-4832

Query Match 29.3%; Score 1345.6; DB 13; Length 1400;
Best Local Similarity 98.6%; Pred. No. 5.3e-293;
Matches 1384; Conservative 5; Mismatches 9; Indels 6; Gaps 3;
3195 AGTGGGAACACAGGCGCTCGCAGCAGCAAGGTGGGGGCAAGCAGAAATGCTCCCAAGGATTT 3254
1 AGTGGGAACACAGGCGCTCGCAGCAGCAAGGTGGGGGCAAGCAGAAATGCTCCCAAGGATTT 60
3255 CACACCTGAGCCCTGCCCCCACCCTGCTGTAAGAAACA-TCCGCCACGTGAAGAGACAGAAG 3313
61 CACACCTGAGCCCTGCCCCCACCCTGCTGTAAGAAACA-TCCGCCACGTGAAGAGACAGAAG 120
3314 GAGGATGGCAGGAGTACCTGGGGAAACAAACAGGAGATCTTTTCTGCCCTGCTCCAG 3373
121 GAGGATGGCAGGAGTACCTGGGGAAACAAACAGG-ATCTTCTGCCCCCTGCTCCAG 176
3374 TCAGATGGGCTGACCCGCTGGATCAGTACCATTTTGGGAGAGAGAGAGAGAGAGAGAGAGAG 3433
177 TCAGATGGGCTGACCCGCTGGATCAGTACCATTTTGGGAGAGAGAGAGAGAGAGAGAGAGAG 236
3434 TTCAGGCTGGGTGAGAGGGGTGGGGAGCCCTTCGGCCCCCTCACCCCTCCAGGCTGCTG 3493
237 TTCAGGCTGGGTGAGAGGGGTGGGGAGCCCTTCGGCCCCCTCACCCCTCCAGGCTGCTG 296
3494 TGAGAGTGTCAAGTGTGTAAGGGGCCCAACCTCAGGTTTCAAGTGTGTAAGGGGCCCAACCTC 3553
297 TGAGAGTGTCAAGTGTGTAAGGGGCCCAACCTCAGGTTTCAAGTGTGTAAGGGGCCCAACCTC 356
3554 TATGCCCGCCGTAGGTAAAGGGGCCCTCTAAACCCCTTGCCCTGCGCTCACCTGGCCAG 3613
357 TATGCCCGCCGTAGGTAAAGGGGCCCTCTAAACCCCTTGCCCTGCGCTCACCTGGCCAG 416
3614 CTACACCCCTTTTGGGTGTAGGGGAAAGAAATGCTGACCCCTGGGAAGGCTCCCTGGGTAGA 3673
417 CTACACCCCTTTTGGGTGTAGGGGAAAGAAATGCTGACCCCTGGGAAGGCTCCCTGGGTAGA 476
3674 ATACACACACTTTTTCAGGTTGTTGCAACACAGGCTCTGAGTTGACCTCTGTTTCAGCCA 3733
477 ATACACACACTTTTTCAGGTTGTTGCAACACAGGCTCTGAGTTGACCTCTGTTTCAGCCA 536
3734 AGGACCAAGAGTGTGTAAGTAAAGTGTCTCAGTCCCCCAGACATGTGCCCTTTGC 3793
537 AGGACCAAGAGTGTGTAAGTAAAGTGTCTCAGTCCCCCAGACATGTGCCCTTTGC 596
3794 TGCTGGCTTACCACTTTTCCCGCAGAGCAGAGGCCCCGAGCCCTTCAGGCCCCAGCACTGC 3853
597 TGCTGGCTTACCACTTTTCCCGCAGAGCAGAGGCCCCGAGCCCTTCAGGCCCCAGCACTGC 656
3854 CCCAGACTCGCTGCACTGCTCCCTCATCTGTAAAGGTGAAGGGGTGATGCAAGATATG 3913
657 CCCAGACTCGCTGCACTGCTCCCTCATCTGTAAAGGTGAAGGGGTGATGCAAGATATG 716
3914 CCTGACAGGAACAGTCTGTGATGGATGATGATGATGATGATGATGATGATGATGATGATGATG 3973
717 CCTGACAGGAACAGTCTGTGATGGATGATGATGATGATGATGATGATGATGATGATGATGATG 776

Qy 3974 CG-TCCGGCGCCCGCAGCCCGCAGCTATCATGTGTCCAGCGTGTGTTTCCCGAGAGCAGAGCT 4032
Db |||||
777 CGCTCCGGCGCCCGCAGCCCGCAGCTATCATGTGTCCAGCGTGTGTTTCCCGAGAGCAGAGCT 836
Qy 4033 CAGCATCACACTGACACTCACCCTCGCCCTGCCCGCAGAGAGGTTACTGCCGAGCGGAC 4092
Db |||||
837 CAGCATCACACTGACACTCACCCTCGCCCTGCCCGCAGAGAGGTTACTGCCGAGCGGAC 896
Qy 4093 TTTGCACTCTGATGACCTCAAAGCATTTCATGCTGCCCTCTGCGAGGGCAGGCGAGGG 4152
Db |||||
897 TTTGCACTCTGATGACCTCAAAGCATTTCATGCTGCCCTCTGCGAGGGCAGGCGAGGG 956
Qy 4153 CAGTGACACTGTAGGAGCATAGCAAGCCAGAGATGGGGTGAAGGGACACAGTCTTGAGC 4212
Db |||||
957 CAGTGACACTGTAGGAGCATAGCAAGCCAGAGATGGGGTGAAGGGACACAGTCTTGAGC 1016
Qy 4213 TGTCCCATCATGTGACTCTCAAACTCTTCCAGATTTCTCTAAGAAATAGCACCCCT 4272
Db |||||
1017 TGTCCCATCATGTGACTCTCAAACTCTTCCAGATTTCTCTAAGAAATAGCACCCCT 1076
Qy 4273 TCCCATTTGCCCGCAGCTTAGCCTCTTCTCCAGGGGAGCTACTCAGGACTCACGTAGCAT 4332
Db |||||
1077 TCCCATTTGCCCGCAGCTTAGCCTCTTCTCCAGGGGAGCTACTCAGGACTCACGTAGCAT 1136
Qy 4333 TAAATCAGCTGTGAATCGTCAGGGGGTGTCTGTAGCCTCAAACCTCTCGGGGCGAGGGAG 4392
Db |||||
1137 TAAATCAGCTGTGAATCGTCAGGGGGTGTCTGTAGCCTCAAACCTCTCGGGGCGAGGGAG 1196
Qy 4393 GCCGAGACTCCGTGGGAGAGCTCATTTCCACATCTTTCGCAAGACAGCTTTGTCCAGCT 4452
Db |||||
1197 GCCGAGACTCCGTGGGAGAGCTCATTTCCACATCTTTCGCAAGACAGCTTTGTCCAGCT 1256
Qy 4453 GTCCACATTCAGTCAGACTGCTCCCGGGGAGAGAGCCCGGCCCGCCAGCATAAAGAAC 4512
Db |||||
1257 GTCCACATTCAGTCAGACTGCTCCCGGGGAGAGAGCCCGGCCCGCCAGCATAAAGAAC 1316
Qy 4513 TGCAAGCTTGGTACTGTCAGAGTCTGGGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTG 4572
Db |||||
1317 TGCAAGCTTGGTACTGTCAGAGTCTGGGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTG 1376
Qy 4573 GGCTGATGACAAATGTTAAAAA 4596
Db |||||
1377 GGCTGATGACAAATGTTAAAAA 1400

RESULT 10
US-60-544-32424
; Sequence 32424, Application US/60680544
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary
; TO INVENTION: Fragments of, Cynomolgus Monkey Genes and the Use Thereof
; FILE REFERENCE: 21590290000
; CURRENT APPLICATION NUMBER: US/60/680,544
; CURRENT FILING DATE: 2005-05-13
; NUMBER OF SEQ ID NOS: 48714
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 32424
; LENGTH: 3658
; TYPE: DNA
; ORGANISM: Macaca Mulatta
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(3658)
; OTHER INFORMATION: n = A,T,C or G
US-60-544-32424

[illegible][illegible]

```
Qy 3185 CCGGCTCA CAGTGGGAAACAGGGCCTCGCAGCAGCAAGGTGGGGGCAAGCAGAAATGCCT 3244
Db 2280 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 3245 CCAGGATTTTCA CACTGAGCCCTGCCCCCTGCTGAAATAAACAATCCGCCAGTGAAG 3304
Db 2340 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 3305 AGACAGAGGAGGATGCGAGGAGTTACCTGGGGAAACAAACAGGGATCTTTTCTGCC 3364
Db 2400 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 3365 CTGCTCAGTCGAGTTGGCCCTGACCCGCTTGGATCAGTGACCAATTTGTTGGCAGACAGG 3424
Db 2460 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 3425 GAGAGCAGCTTCAGCCTGGGTGAGAAGGGGTGGCGAGCCCTTCGGCCCTCACCTCC 3484
Db 2520 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 3485 AGGCTGCTGTGAGAGTGTCAAGTGTGTAGGGGCCAAACTCAGGTTCAAGTGCAAAACAG 3544
Db 2580 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 3545 GTCAGCAGGTATGCCGCCCGCTAGGTTAAGGGGCCCTCTAAACCCCTTCGCTGGCCTCA 3604
Db 2640 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 3605 CTTGGCCAGCTCACCCCTTTGGGTGAGGGGAAAGAAATGCCCTGACCCTGGGAAGGCTC 3664
Db 2700 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 3665 CTTGGTAGAATACACCACACTTTTTCAGGTTGTTGCAACA CAGGTCCTGAGTGCACCTCG 3724
Db 2760 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 3725 GTTCAGCCAAAGCAAGAGGTGTGTAAGTGAAGTGGTTCTCAGTCCCGCAGACATGTG 3784
Db 2820 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 3785 CCCTTTGCTGGCTACACTCTTCCCGCAGACGACGCGCCCGCAGCCCTTCAGGCC 3844
Db 2880 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 3845 CAGCACTGCCCAGACTCGCTGGCACTCAGTTCCTCATCTGTAAAGGTGAAGGTGATG 3904
Db 2940 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 3905 CAGGATATGCTGACAGAAACAGTCTGTGGATGGACATGATGCTTAAGGAAAGCAGC 3964
Db 3000 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGA
Qy 3965 AGAGAGAGAGCTCGGGCGCCCGACCCCACTACAGTGTCCAGGTGCTGGTCCCCAGA 4024
Db 3060 CGC-----TCGGCAGCCCCAGCCCACTGTGAGGCTCAGCGTGGGGGCCCCCAGCA 3111
Qy 4025 GCACAGCTCAGCATCAGCTGACACTCACCTGCCCTGCCCTGGCCAGAGGGTACTGCC 4084
Db 3112 GCACAGC-----TCACAAAAACACTCACCTGCCCTGCCCGGGCGGAGGGTACTGCT 3165
Qy 4085 GAGCGCACTTTTGCACTCTGATGACCTCAAGACACTTTTCATGGGTGCCCTCTGGCAGGCA 4144
Db 3166 GACAGCACTTTTGCACTCTGATGACCTCAAGACACTTTTCATGGGTGCCCTCT-----GGCA 3220
Qy 4145 GGCAGGGCAGTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGACACAG 4204
Db 3221 GGCAGGTCAGTGACTCTGTAGGAGCATAGCAGTCCAGGAGATGGGGTGAAGG--ACAG 3278
Qy 4205 TCTTGAGCTGTCCACATGCGATGTGACTCTCTCAAACTCTTCCAGATTTCTCTAAGATAG 4264
Db 3279 TCTTGAGCTGTCTACATGTCATGTGACTCTCTCAAACTCTCTCCAGATTTCTCT---AATAG 3335
```

```
Qy 4265 CACCCCTTTCCCATTTGCCCCAGCTTAGCCTCTTCTCCAGGGAGGAGCTACTCAGGACTCA 4324
Db 3336 CAACCACTTTCCCATTTGCCCCAGCTTAGCCTCTTCTCCAGGGAGGAGCTACTCAGGACTCA 3395
Qy 4325 CGTAGCATTAATAACAGCTGTGAATCGTCA--GGGGGTCTCTGTAGCCTCAACCTCTCTGGG 4383
Db 3396 TGTAGCATTAAGTGTGAGTGTGAAATTTGTGAGGGGGGTGTGAGTGTCTCAACCTCTCTGGG 3455
Qy 4384 GCAGGGGACCCCGAGACTCTCGTGGGAGAACTCATTTCCACATCTTTGCCAAGACAGCCTT 4443
Db 3456 GCGGGGAGCCCGAGACTCTCTGGGAGAACTCATCCCCAGCTTGTG--NNNNACAGCCTT 3514
Qy 4444 TGTCCAGCTGTCCACATTTGAGTGTGAGTGTCTCCCGGGGAGAGAGCCCCGGCCCCCAGCAC 4503
Db 3515 TGTCCAGCTGTCCACAGCGAGCAGACTGTCTCCGGGGAGACAGCCCCGGCCCCCAGCAC 3574
Qy 4504 ATAAAGAACTGACAGCTTGTGACTGTGAGACTCTGGGTTGTAGAGAACTCTTTGTAAGCAA 4563
Db 3575 ATAAAGAACTGACAGCTTGTGACTGTG--AGTGTGGGTTGTAGAGAACTCTTTGTAAGCAA 3632
Qy 4564 TAAAGTTTGGGGTGATGACAAATGTT 4589
Db 3633 TAAAGTTTGGGGTGATGACAGATGTT 3658
```

RESULT 12

```
US-10-450-763-15296
; Sequence 15296, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 15296
; LENGTH: 722
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (157)..(468)
; OTHER INFORMATION: 68% homologous to Mus musculus Pro-Pol-dUTPase
; OTHER INFORMATION: polyprotein, accession number Y12713, Smith-Waterman Score=354.
US-10-450-763-15296
```

```
Query Match 13.7%; Score 627.6; DB 8; Length 722;
Best Local Similarity 94.8%; Pred. No. 2.6e-131;
Matches 681; Conservative 0; Mismatches 34; Indels 3; Gaps 3;

Qy 2564 TCCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGCAGGAGCAATTTCTCTCGTGCC 2623
Db 5 TTCTTAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGCAGGAGCAATTTTCTCTCGTGCC 64

Qy 2624 TCAGCATTCAGACAGCCTCTCCCTCTCGGATGACAGTGAAGAGAAACCCATCAAGGCTCTC 2683
Db 65 TCAGCATTCAGACAGCCTCTCCCTCTCGGATGACAGTGAAGAGAAACCCATCAAGGCTCTC 124

Qy 2684 AAAGCTCGGGGACACCTTGTAGCTCAGGCGTACACTCTCGAGCAGCAGCCGAGGCTC 2743
Db 125 AAAGCTCGGGGACACCTTGTAGCTCAGGCGTACACTCTCGAGCAGCAGCCGAGGCTC 184

Qy 2744 GAAGCTCCAGCTGGAACATGGTGTGCTGGCCCGGGGGCGGCCACCGACACCCAGCTATT 2803
Db 185 GAAGCTCCAGCTGGAACATGGTGTGCTGGCCCGGGGGCGGCCACCGACACCCAGCTATT 244
```

QY 2804 TCAATGGTGTGAAGTCCAAATACAGTCTCTTTAAATGGTGAAACACCTGCACATCCGGAGT 2863
Db |||||||
QY 245 TCAATGGTGTGAAGTCCAAATACAGTCTCTTTAAATGGTGAAACACCTGCACATCCGGAGT 304
Db |||||||
QY 2864 TCCACCGGTCAAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCGAGTGCAG 2923
Db |||||||
QY 305 TCCACCGGTCAAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCGAGTGCAG 364
Db |||||||
QY 2924 CTTTCAGCTTGTGTACCAAGAGCGGCAGCCTGTTGCTACACATGAGGTGCCAGACT 2983
Db |||||||
QY 365 CTTTCAGCTTGTGTACCAAGAGCGGCAGCCTGTTGCTACACATGAGGTGCCAGACT 424
Db |||||||
QY 2984 CGGGCATGCACCTGCAGTGCACACTGGCCCTGATGGCAGCTTTCCTGGAGCTGGAGGG 3043
Db |||||||
QY 425 CGGGCATGCACCTGCAGTGCACACTGGCCCTGATGGCAGCTTTCCTGGAGCTGGAGGG 484
Db |||||||
QY 3044 TCAAGCATGGCCAGCTGGAGAACAGGCCCTAAACCTGCTCCACCGCGGCTCCACACT 3103
Db |||||||
QY 485 TCAAGCATGGCCAGCTGGAGAACAGGCCCTAAACCTGCTCCACCGCGGCTCCACACT 544
Db |||||||
QY 3104 GCGGAAAGCAGCCTTCTGTC-TGGTGACAGATGC-TGCCCTGAAACACAGGCTCAGC 3161
Db |||||||
QY 545 TGCCGGAAGCAGCCTTCTGCTTCCGTGCAAGATGCTTGCCTTGAACACAGGCTCAGC 604
Db |||||||
QY 3162 CGTTCCAGGGGAT-TGCCAGGCCCCCGGCTCACAGTGGGAACAGGCGCTCGCAGCAGC 3220
Db |||||||
QY 605 CGTTCCAGGGGATCTTCCAGCCCCCGGCTCAGCAGTGGGACAGGCGCTCGCAGCAGC 664
Db |||||||
QY 3221 AAGTGGGGGCAAGCAGAAATGCTCCAGGATTTTCAACCTGAGCCCTGCCACCCCT 3278
Db |||||||
QY 665 AAGTGGGGGCAAGCAGAAAGCTCCCAAGATTTTACAACCTGGGCGCTTGGCCCCCT 722
Db |||||||

RESULT 13
US-60-680-544-10785/c
; Sequence 10785, Application US/60680544
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary
; FILE REFERENCE: 21590290000
; CURRENT APPLICATION NUMBER: US/60/680,544
; CURRENT FILING DATE: 2005-05-13
; NUMBER OF SEQ ID NOS: 48714
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 10785
; LENGTH: 761
; TYPE: DNA
; ORGANISM: Macaca Fascicularis
; NAME/KEY: misc_feature
; LOCATION: (1)...(761)
; OTHER INFORMATION: n = A,T,C or G
US-60-680-544-10785

Query Match 9.6%; Score 440.6; DB 15; Length 761;
Best Local Similarity 72.6%; Pred. No. 3.6e-89;
Matches 566; Conservative 0; Mismatches 192; Indels 22; Gaps 8;

QY 3777 GACATGTGCCCTTTGCTGTGGTACCACCTCTTCCCGAGCAGCAGGCCCGCCGAGCCCC 3836
Db |||||||
QY 761 GACACATGNNNNNNNGCTGTGGTATCATNNNTTCCCGAGCGCAGCAGNNNNNGAGCCCC 702
Db |||||||
QY 3837 TTCCAGGCCAGCAGTGCAGCAGCTCGTGGCAGCTCAGTTCCTCATCTGTAAGGTGAA 3896
Db |||||||
QY 701 TTCAGGCCAG- -NN 644
Db |||||||

QY 3897 GGGTCATCGAGATATGCTGCAGGAAACAGTCTGTGGATGGAATGATCACTCACTAAGG 3956
Db |||||||
QY 643 NNN 584
Db |||||||
QY 3957 AAAGCAGCAGAGAGAGAGCG-TCCGGCGCCCGAGCCCACTATCAGTGTCCAGCGTGTGG 4015
Db |||||||
QY 583 AAAGCAGCAGAGAGAGAGCGTCCGGCACCAGCCCACTGTGACGCTCCAGCGTGTGGGG 524
Db |||||||
QY 4016 TTCCCGCAGCAGCAGCTCAGCATCACTGACACTCACCTGCTGCCCTGCCCTGGCCAGAG 4075
Db |||||||
QY 523 CCCCCCAGCAGCAGC-----TCACACGAACTCACCTGCTGCCCGCCCGCGCGGAG 470
Db |||||||
QY 4076 GGTACTCGCAGCGGAGCTTTGCACTCTGATGACCTCAAAGCAGCTTTCATGGCTGCCCTCT 4135
Db |||||||
QY 469 GGTANNN 410
Db |||||||
QY 4136 GCGAGGCGAGCGGCGAGTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGA 4195
Db |||||||
QY 409 -----GGCAGGCGAGTCACTGTAGGAGC--ANNNNNNNNNNNNNNNNNNNNNNNN 357
Db |||||||
QY 4196 GGGACACAGTCTTGAGCTGTCCACATGCATGTGACTCTCTCAAACCTCTTCAGATTCTC 4255
Db |||||||
QY 356 NNN 297
Db |||||||
QY 4256 TAAGAAATAGCAGCCCTTCCCATTTGCCAGCTTAGCTCTTCCCGAGGGAGCTACT 4315
Db |||||||
QY 296 T---AATAGCAACACTTCCCCCATTTGCCCTTAGCTCTTCTCCAGGGAGCTACT 240
Db |||||||
QY 4316 CAGGACTCACCTAGCATTAAATCAGCTGTAATGCTCA-GGGGGTGTCTGTAGCTCAA 4374
Db |||||||
QY 239 CAGACTCATGTAGCATTAAGTCACTGTGATTTGTGAGGGGGTGTGAGCTCA 180
Db |||||||
QY 4375 CCTCTGGGCGAGGGGAGCGCAGACTCCGTGGGAGAGCTCATTTCCACATCTTGCCAA 4434
Db |||||||
QY 179 CCTCTGGGCGGGGAGCGCGAGACTCTTTGGGAGAGCTCATCCCCACCTCTTGCCAA 120
Db |||||||
QY 4435 GACAGCCTTTGTCAGCTGTCCACATTTAGTCACTGCTCCCGGGGAGAGCCCCCGC 4494
Db |||||||
QY 119 GACAGCCTTTGTCAGCTGTCCACAGCGAGCAGACTGCTCCCGGGGAGAGAGCCCCCGC 60
Db |||||||
QY 4495 CCCCAGCACATAAAGAACTGCGAGCTTGGTACTGACAGAGTCTGGGTTGTAGAGAACTCTT 4554
Db |||||||
QY 59 CCCCAGCACATAAAGAACTGCGAGCTTGGTACTGCG--AGTGTGGGTTGTAGAGAACTCTT 2
Db |||||||

RESULT 14
US-60-680-544-11056
; Sequence 11056, Application US/60680544
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary
; FILE REFERENCE: 21590290000
; CURRENT APPLICATION NUMBER: US/60/680,544
; CURRENT FILING DATE: 2005-05-13
; NUMBER OF SEQ ID NOS: 48714
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 11056
; LENGTH: 761
; TYPE: DNA
; ORGANISM: Macaca Fascicularis
; NAME/KEY: misc_feature
; LOCATION: (1)...(761)
; OTHER INFORMATION: n = A,T,C or G

Job time : 1775.13 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:40:00 ; Search time 144.36 Seconds
(without alignments)
2537.148 Million cell updates/sec

Title: US-09-155-676B-7

Perfect score: 5052
Sequence: 1 MAVNMACPGAGSGAVGQK.....PDGSFAWSRVKRGHLENRP 947

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	5052	100.0	947	2	AAW42402 Amino aci
2	5052	100.0	947	5	AAU80371 Human cel
3	5052	100.0	947	7	ADP45075 Human kin
4	5052	100.0	947	8	ADS88253 Human pro
5	5045	99.9	947	2	AAI31665
6	5044	99.8	947	2	AAW82497 Human NIK
7	5044	99.8	947	2	AAW81564 Human NF-
8	5044	99.8	947	2	AAW81561 Human NF-
9	5040	99.8	947	7	AAI31667
10	5040	99.8	947	7	ABM85491 Human pro
11	5032	99.6	947	2	AAI31668
12	3439.5	68.1	697	8	ADP29813 Human sec
13	2481.5	49.1	548	7	ABM85490 Mouse pro
14	2301	45.5	477	8	ADQ67717 Novel hum
15	1703	33.7	324	7	ABR61616 Human NIK
16	1703	33.7	324	7	ABR61608 Human NIK
17	1101	21.8	213	2	AAI31666
18	479	9.5	94	3	AAQ02579 Human sec
19	446	8.8	81	7	ABR61615 Human NIK
20	446	8.8	81	7	ABR61607 Human NIK
21	367	7.3	250	5	ABB78798 Serine/th
22	367	7.3	256	5	ADI11263 Polypepti
23	367	7.3	256	5	ADI117346 Polypepti
24	367	7.3	256	5	ADI117320 Polypepti
25	367	7.3	256	5	ADI117340 Polypepti

RESULT 1
AAW42402
ID AAW42402 standard; protein; 947 AA.

XX AC AAW42402;

XX DT 15-APR-1998 (first entry)

XX DE Amino acid sequence of NF-kappaB inducing kinase.

XX KW Human tumour necrosis factor receptor-associated factor 2; TRAF2;

KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;

KW intracellular signalling activity; acute hepatitis;

KW autoimmune-induced cell death.

XX OS Homo sapiens.

XX PN W09737016-A1.

XX PD 09-OCT-1997.

XX PF 01-APR-1997; 97WO-IL000117.

XX PR 02-APR-1996; 96IL-00117800.

XX PR 26-AUG-1996; 96IL-00119133.

XX PA (YEDA) YEDA RES & DEV CO LTD.

XX PI Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;

XX DR WPI; 1997-503101/46.

XX DR N-PSDB; AAV03325.

XX PT DNA encoding tumour necrosis factor receptor-associated factor binding

XX PT molecule - used for modulation or mediation in cells of the activity of

XX PS NF-kB.

XX PS Disclosure; Fig 6; 127pp; English.

XX CC The present sequence represents a NF-kappa inducing kinase (NIK). The

XX CC full length DNA sequence was obtained by PCR using clone 10 (AAV03324).

XX CC NIK was found to induce NF-kappaB even more effectively than TRAF2.

XX CC Proteins capable of binding to TRAF2 and NIK were identified. The NIK or

XX CC TRAF-2 binding proteins can be used for modulation or mediation in cells

XX CC of NF-kappaB activity or any other intracellular signalling activity

XX CC modulated or mediated by TRAF2. TRAF-binding proteins are especially used

XX CC for prevention or treatment of pathological conditions associated with NF

XX CC -kappaB induction, e.g. acute hepatitis, autoimmune-induced cell death,

CC e.g. death of the beta Langerhans cells or the pancreas that results in
CC diabetes, the death of cells in graft rejection, the death of
CC oligodendrocytes in the brain in multiple sclerosis, and AIDS-inhibited T
CC cell suicide which causes proliferation of the AIDS virus and hence the
CC AIDS disease. The proteins are also useful for screening of ligands
CC capable of binding to a protein, which are useful for modulating cellular
CC activity modulated/mediated by TRAF2
XX
SQ Sequence 947 AA;

Query Match 100.0%; Score 5052; DB 2; Length 947;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVMEMACPGAGSAGVGOQKELPKPKETPPGLKQSSVYKLEAVEKSPVFCGKWEILND 60
DB 1 MAVMEMACPGAGSAGVGOQKELPKPKETPPGLKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKGSEAGPAAISIIAQAECEQSQFSPFSEIFIAAGSKQYSQSESQDQIPNN 120
DB 61 VITKGTAKGSEAGPAAISIIAQAECEQSQFSPFSEIFIAAGSKQYSQSESQDQIPNN 120
QY 121 VAHATEGKMARVCWKGRSKARKKRRKSKSLAHAGVALAKPLPTPEQESCTIPVOE 180
DB 121 VAHATEGKMARVCWKGRSKARKKRRKSKSLAHAGVALAKPLPTPEQESCTIPVOE 180
QY 181 DESPLGAPVYRNTPOFTKPKGPGQLCFKQGLRGLRPAALPRSELHKLISPLQCLNHVW 240
DB 181 DESPLGAPVYRNTPOFTKPKGPGQLCFKQGLRGLRPAALPRSELHKLISPLQCLNHVW 240
QY 241 KLHHPODGGPLPLTHFFPYSLPHPPFPHPQPKPPLSFLGKLACVDSQKPLDPDH 300
DB 241 KLHHPODGGPLPLTHFFPYSLPHPPFPHPQPKPPLSFLGKLACVDSQKPLDPDH 300
QY 301 LSKLACVDSQKPLPGHLEPSCLSRGAHEKFSVEYLVHALQGSVSSQAHSLSLAKTW 360
DB 301 LSKLACVDSQKPLPGHLEPSCLSRGAHEKFSVEYLVHALQGSVSSQAHSLSLAKTW 360
QY 361 AARGSRSPSKPTEDNEGVLTATEKLKPDVYREVEVHWATHQLRGRSGFGEVHRMEDK 420
DB 361 AARGSRSPSKPTEDNEGVLTATEKLKPDVYREVEVHWATHQLRGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVPLYGAVREGPWNIIFWELLEGSLGQL 480
DB 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVPLYGAVREGPWNIIFWELLEGSLGQL 480
QY 481 VKEQCLPEDRALYILGQALEGLYHLSRIILHGVNDKADNVLSSDGSAAALCDFGHAVC 540
DB 481 VKEQCLPEDRALYILGQALEGLYHLSRIILHGVNDKADNVLSSDGSAAALCDFGHAVC 540
QY 541 LOPDGLGKSLTGDIYPTGTHMAPEVVLGRSCDAKVDVWSSCCMWLHMGCHPWTQFF 600
DB 541 LOPDGLGKSLTGDIYPTGTHMAPEVVLGRSCDAKVDVWSSCCMWLHMGCHPWTQFF 600
QY 601 RGPLCLKIASEPPVREIPPPSCAPLTAQAIQELRKEPIHRVSAALGKKNRALQQVGG 660
DB 601 RGPLCLKIASEPPVREIPPPSCAPLTAQAIQELRKEPIHRVSAALGKKNRALQQVGG 660
QY 661 LKSPWRGEYKPRHPHPNQAQYHQTTHAQPRELSPRAPCPRAEETTGAPKLQPLPPE 720
DB 661 LKSPWRGEYKPRHPHPNQAQYHQTTHAQPRELSPRAPCPRAEETTGAPKLQPLPPE 720
QY 721 PPEPNKSPPLTUSKEGSGWPEPLPLSSLEPAPARNPSSPERKATVPQEOLQLELEFLN 780
DB 721 PPEPNKSPPLTUSKEGSGWPEPLPLSSLEPAPARNPSSPERKATVPQEOLQLELEFLN 780
QY 781 SLSQFSLSEEQILSCLSIDSLSDSEKNPASKASQSSRDTLSSGVHSSSQAEARSS 840
DB 781 SLSQFSLSEEQILSCLSIDSLSDSEKNPASKASQSSRDTLSSGVHSSSQAEARSS 840
QY 841 SNMVLARGRPDTPSYNGVKVQIQLNGEHLHIREFHRVKVGDIAATGSIQIPAAAFS 900
DB 841 SNMVLARGRPDTPSYNGVKVQIQLNGEHLHIREFHRVKVGDIAATGSIQIPAAAFS 900

QY 901 LVTKDQGPVRYDMEVPSDGSIDLOCTLAPDGSFAMSRVVKHGOLENRP 947
DB 901 LVTKDQGPVRYDMEVPSDGSIDLOCTLAPDGSFAMSRVVKHGOLENRP 947

RESULT 2

AAU80371
ID AAU80371 standard; protein; 947 AA.

XX AC AAU80371;

XX DT 30-JUL-2002 (first entry)

XX DE Human cellular kinase NIK protein.

XX KW Human; virucide; cytomegalovirus infection; CMV; cellular kinase; RICK;
RIP; Nck-Interacting kinase; NIK; MKK3; SRPK-2.

XX OS Homo sapiens.

XX PN EPI201765-A2.

XX PD 02-MAY-2002.

XX PF 15-OCT-2001; 2001EP-00124604.

XX PR 16-OCT-2000; 2000US-0240750P.

XX PA (AXXI-) AXKIMA PHARM AG.

XX PI Schubart D, Habenberger P, Stein-Gerlach M, Bevec D;

XX DR WPI; 2002-373930/41.

XX DR N-PSDB; ABK51171.

PT Identifying agents for treatment or prevention of cytomegalovirus
PT infection, comprises contacting test compound with cellular kinase and
PT detecting change in cellular kinase activity.

XX PS Disclosure; Page 34-36; 49pp; English.

CC The present invention relates to a new method for identifying compounds
CC for treating and/or preventing cytomegalovirus (CMV) infection and/or
CC related diseases. The method of the invention comprises contacting a test
CC compound with at least one of the cellular kinases RICK, RIP, Nck-
CC interacting kinase, MKK3 and SRPK-2 and detecting any change in kinase
CC activity. The method of the invention can be used to treat and/or prevent
CC CMV infections and related diseases. Oligonucleotides that can detect the
CC specified kinases can also be used for diagnosis of infection. The
CC present amino acid sequence represents the human cellular kinase NIK (Nck
CC -interacting kinase) protein of the invention, as described above

SQ Sequence 947 AA;

Query Match 100.0%; Score 5052; DB 5; Length 947;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVMEMACPGAGSAGVGOQKELPKPKETPPGLKQSSVYKLEAVEKSPVFCGKWEILND 60
DB 1 MAVMEMACPGAGSAGVGOQKELPKPKETPPGLKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKGSEAGPAAISIIAQAECEQSQFSPFSEIFIAAGSKQYSQSESQDQIPNN 120
DB 61 VITKGTAKGSEAGPAAISIIAQAECEQSQFSPFSEIFIAAGSKQYSQSESQDQIPNN 120
QY 121 VAHATEGKMARVCWKGRSKARKKRRKSKSLAHAGVALAKPLPTPEQESCTIPVOE 180
DB 121 VAHATEGKMARVCWKGRSKARKKRRKSKSLAHAGVALAKPLPTPEQESCTIPVOE 180
QY 181 DESPLGAPVYRNTPOFTKPKGPGQLCFKQGLRGLRPAALPRSELHKLISPLQCLNHVW 240

Db 181 DESPLGAPYVNTQFTKPLKEBPGQLGQLCFKQLEGLRALPRSELHKLISPLQCLNHW 240
Qy 241 KLHPDQGGPLPLTPHPFPYSRLEPHFPPLQPMKPHPLESFLGKLACVDSQKPLDPDH 300
Db 241 KLHPDQGGPLPLTPHPFPYSRLEPHFPPLQPMKPHPLESFLGKLACVDSQKPLDPDH 300
Qy 301 LSKLACVDSPKPLPGPHLEPSCLSRGAHEKFSVEYLVLHVALQGSVSSQAHSLSLAKTW 360
Db 301 LSKLACVDSPKPLPGPHLEPSCLSRGAHEKFSVEYLVLHVALQGSVSSQAHSLSLAKTW 360
Qy 361 AAGRSRSPKPTKEDNEGVLITLTKLPVDYVEEVEVHWATHQLRLGRSGFGEVHRMEDK 420
Db 361 AAGRSRSPKPTKEDNEGVLITLTKLPVDYVEEVEVHWATHQLRLGRSGFGEVHRMEDK 420
Qy 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIVLYGAVREGPWNIIFMELLEGSLGQL 480
Db 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIVLYGAVREGPWNIIFMELLEGSLGQL 480
Qy 481 VKEQGCLPEDRALYVLGQALEGLYLSRRILHGDVADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGCLPEDRALYVLGQALEGLYLSRRILHGDVADNVLLSSDGSAAALCDFGHAVC 540
Qy 541 LQPDGLKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Db 541 LQPDGLKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Qy 601 RGPLCLKIASEPPPPVREIPSPCAPLTAQAIQEGRLKEPIHRVSAELGGKVNRLAQVGG 660
Db 601 RGPLCLKIASEPPPPVREIPSPCAPLTAQAIQEGRLKEPIHRVSAELGGKVNRLAQVGG 660
Qy 661 LKSPWRGEYKEPRHPNPNQANYHOTLHAQPRELSRAPGPAPBETTTGRAPKLPPLPPE 720
Db 661 LKSPWRGEYKEPRHPNPNQANYHOTLHAQPRELSRAPGPAPBETTTGRAPKLPPLPPE 720
Qy 721 PPEPNKSPPLTSLKEESGMWELPLSLLEPAPARNPSPERKATVPQELQOLEIEFLN 780
Db 721 PPEPNKSPPLTSLKEESGMWELPLSLLEPAPARNPSPERKATVPQELQOLEIEFLN 780
Qy 781 SLSQPFSLKEEQEILSCLSIDSLSDSEKNPSKASOSSRDTLSSGVHSSQAARSS 840
Db 781 SLSQPFSLKEEQEILSCLSIDSLSDSEKNPSKASOSSRDTLSSGVHSSQAARSS 840
Qy 841 SNNVLARGRTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGISSQIPAAAFS 900
Db 841 SNNVLARGRTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGISSQIPAAAFS 900
Qy 901 LVTKDGPVRYDMEVPDSDIGDLOCTLAPDGSFAWSRWKHCQLENRP 947
Db 901 LVTKDGPVRYDMEVPDSDIGDLOCTLAPDGSFAWSRWKHCQLENRP 947

RESULT 3

ADP45075
ID ADP45075 standard; protein; 947 AA.

XX AC ADP45075;

XX DT 12-FEB-2004 (first entry)

XX DE Human kinase NIK.

XX KW Human; protein kinase; enzyme; inhibitor; NIK.

XX OS Homo sapiens.

XX PN WO2003081210-A2.

XX PD 02-OCT-2003.

XX PF 20-MAR-2003; 2003WO-US008725.

XX PR 21-MAR-2002; 2002US-0366892P.

XX

PA (SUNE-) SUNESIS PHARM INC.

PI Prescott JC, Braisted A;

XX WPI; 2003-865136/80.

XX Identifying ligand binding to inactive conformation of target protein
kinase (T) comprises contacting the conformation modified (T) which
contains reactive group at binding site, with ligands and detecting
kinase-ligand conjugate formation.

XX Disclosure; SEQ ID NO 44; 260pp; English.

XX The present invention relates to a method for identifying a ligand (L),
which binds to an inactive conformation of target protein kinase (T). The
method involves contacting inactive conformation of (T), which contains
or is modified to contain a reactive group at or near a binding site of
interest, with one or more ligand candidates capable of covalently
bonding to the reactive group thus forming a kinase-(L) conjugate (C).
The method is useful for identifying protein kinase inhibitors that
preferentially bind to inactive conformation of a target protein kinase.
The present sequence is a protein kinase which may be modified via an
amino acid substitution, for use in the method of the invention.

XX Sequence 947 AA;

Query Match 100.0%; Score 5052; DB 7; Length 947;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAVNEMACPGAGSAGVQKQKELPKPKTKTTPGLKQKQSVYKLEAVEKSPVFCGKWEILND 60

Db 1 MAVNEMACPGAGSAGVQKQKELPKPKTKTTPGLKQKQSVYKLEAVEKSPVFCGKWEILND 60

Qy 61 VITKGTAKESAGPAAISIIIAQACENSQSFSTFSEIRIFIAQSKQVSESLEQIINN 120

Db 61 VITKGTAKESAGPAAISIIIAQACENSQSFSTFSEIRIFIAQSKQVSESLEQIINN 120

Qy 121 VAHATEGMARVCKWKGKRRSKARKKRRKSKSLAHAGVALAKPLPRTPPEQESCTIPVQE 180

Db 121 VAHATEGMARVCKWKGKRRSKARKKRRKSKSLAHAGVALAKPLPRTPPEQESCTIPVQE 180

Qy 181 DESPLGAPYVNTQFTKPLKEBPGQLGQLCFKQLEGLRALPRSELHKLISPLQCLNHW 240

Db 181 DESPLGAPYVNTQFTKPLKEBPGQLGQLCFKQLEGLRALPRSELHKLISPLQCLNHW 240

Qy 241 KLHPDQGGPLPLTPHPFPYSRLEPHFPPLQPMKPHPLESFLGKLACVDSQKPLDPDH 300

Db 241 KLHPDQGGPLPLTPHPFPYSRLEPHFPPLQPMKPHPLESFLGKLACVDSQKPLDPDH 300

Qy 301 LSKLACVDSPKPLPGPHLEPSCLSRGAHEKFSVEYLVLHVALQGSVSSQAHSLSLAKTW 360

Db 301 LSKLACVDSPKPLPGPHLEPSCLSRGAHEKFSVEYLVLHVALQGSVSSQAHSLSLAKTW 360

Qy 361 AAGRSRSPKPTKEDNEGVLITLTKLPVDYVEEVEVHWATHQLRLGRSGFGEVHRMEDK 420

Db 361 AAGRSRSPKPTKEDNEGVLITLTKLPVDYVEEVEVHWATHQLRLGRSGFGEVHRMEDK 420

Qy 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIVLYGAVREGPWNIIFMELLEGSLGQL 480

Db 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIVLYGAVREGPWNIIFMELLEGSLGQL 480

Qy 481 VKEQGCLPEDRALYVLGQALEGLYLSRRILHGDVADNVLLSSDGSAAALCDFGHAVC 540

Db 481 VKEQGCLPEDRALYVLGQALEGLYLSRRILHGDVADNVLLSSDGSAAALCDFGHAVC 540

Qy 541 LQPDGLKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600

Db 541 LQPDGLKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600

Qy 601 RGPLCLKIASEPPPPVREIPSPCAPLTAQAIQEGRLKEPIHRVSAELGGKVNRLAQVGG 660

Db 601 RGPLCLKIASEPPPPVREIPSPCAPLTAQAIQEGRLKEPIHRVSAELGGKVNRLAQVGG 660

QY 841 SWNVLAGRRTDTPSYFNGVKVQIQLNGEHLHIREFHRVKGVDIATGISSQIPAAAFS 900
 DB 841 SWNVLAGRRTDTPSYFNGVKVQIQLNGEHLHIREFHRVKGVDIATGISSQIPAAAFS 900
 QY 901 LVTXGQGVRYDMEVPSGIDLOCTLAPDGSFAMSVKVGQLENRP 947
 DB 901 LVTXGQGVRYDMEVPSGIDLOCTLAPDGSFAMSVKVGQLENRP 947

RESULT 5
 AAY31665
 ID AAY31665 standard; protein; 947 AA.
 AC AAY31665;
 XX
 XX 09-NOV-1999 (first entry)
 XX NF-kB inducing kinase (NIK).
 XX
 XX NIK; NF-kB inducing kinase; inflammation; tumour necrosis factor;
 KW interleukin-1; cytokine; inhibitor; antiinflammatory; apoptosis;
 KW anti-apoptotic; human.
 XX
 XX Homo sapiens.
 XX
 XX WO9943704-A1.
 XX
 XX 02-SEP-1999.
 XX
 XX 25-FEB-1999; 99WO-US004110.
 XX
 XX 27-FEB-1998; 98US-0076299P.
 XX (REGC) UNIV CALIFORNIA.
 XX
 XX Greene WC, Lin X, Gelezuinas R;
 XX WPI; 1999-518837/43.
 DR N-PSDB; AAX87842.
 XX
 XX New NF-kB inducing kinase proteins, used for inhibiting a kB-dependent
 PT immune response, e.g. an inflammatory response or an anti-apoptotic
 PT response.
 XX
 XX Disclosure; Fig 10; 48pp; English.
 XX
 CC The present sequence represents human wild-type NIK, a NF-kB inducing
 CC kinase having serine/threonine kinase activity. The invention relates to:
 CC (a) an N-terminus deletion mutant NIK protein (see AAY31666); and (b) a
 CC kinase deficient NIK mutant protein (see AAY31667 and AAY31668) that
 CC inhibits auto-phosphorylation or transphosphorylation. The invention
 CC provides the molecular basis for cytokine induction of NF-kB-dependent
 CC immune and inflammatory responses, emphasising a role for both NIK-NIK
 CC and NIK-IKK (IKK-specific kinase) interactions. A novel and highly
 CC specific method for modulating NF-kB-dependent immune, inflammatory and
 CC anti-apoptotic responses is based on interruption of the critical protein
 CC interaction of NIK and IKK. The mutant NIK proteins are used in claimed
 CC methods for inhibiting NF-kB-dependent gene expression. The kinase-
 CC deficient NIK mutant proteins inhibit activation of IKK. The N-terminal
 CC deletion NIK mutant proteins bind to IKK, thus inhibiting NIK/IKK
 CC interaction
 XX
 XX Sequence 947 AA;
 SQ

Query Match 99.9%; Score 5045; DB 2; Length 947;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 946; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MAVNEMACPGAPGAVGQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
 DB 1 MAVNEMACPGAPGAVGQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
 QY 61 VITKGTAKGSEAGPAAISIIAQACENSQSFSTFSERIFIAGSKOYSQSESIDQIPNN 120

DB 61 VITKGTAKGSEAGPAAISIIAQACENSQSFSTFSERIFIAGSKOYSQSESIDQIPNN 120
 QY 121 VAHATEGKMARVCKWKRRSKARKKRRKSKSLAHAGVALAKPLRTPPEOSTCTIPVOE 180
 DB 121 VAHATEGKMARVCKWKRRSKARKKRRKSKSLAHAGVALAKPLRTPPEOSTCTIPVOE 180
 QY 181 DESPLGAPYVRNTPQFTKPLKEPGLQCFKQKQEGELRPAALPRSELHKLISPLQCLNHVW 240
 DB 181 DESPLGAPYVRNTPQFTKPLKEPGLQCFKQKQEGELRPAALPRSELHKLISPLQCLNHVW 240
 QY 241 KLHHPQDGGPLPLTPHPFPYSRLPHPPHPLQDQKPHPLESFLGKLACVDQSOKPLPDPH 300
 DB 241 KLHHPQDGGPLPLTPHPFPYSRLPHPPHPLQDQKPHPLESFLGKLACVDQSOKPLPDPH 300
 QY 301 LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVAHALQGSVSSQAHSLSLAKTW 360
 DB 301 LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVAHALQGSVSSQAHSLSLAKTW 360
 QY 361 AARGSRREPSPKTEDNEGVLITKLPDYVEYEEVHWATHQLRGRGSGFGEVHRMEDK 420
 DB 361 AARGSRREPSPKTEDNEGVLITKLPDYVEYEEVHWATHQLRGRGSGFGEVHRMEDK 420
 QY 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIPLYGAVREGPWNI FMELEGGSLQOL 480
 DB 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIPLYGAVREGPWNI FMELEGGSLQOL 480
 QY 481 VKEQGCLPEDRALYILGOALEGLEYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
 DB 481 VKEQGCLPEDRALYILGOALEGLEYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
 QY 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVMSSCCMMLHMLNGCHPWTQFF 600
 DB 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVMSSCCMMLHMLNGCHPWTQFF 600
 QY 601 RGPLCLKIASPPPPVREIPPCAPLTAQAIOEGLRKEPIHRVSAEELGGKVNRAALQQVGG 660
 DB 601 RGPLCLKIASPPPPVREIPPCAPLTAQAIOEGLRKEPIHRVSAEELGGKVNRAALQQVGG 660
 QY 661 LKSPWRGEYKPPHPPPNQANYHOTLHAQPRELSPRAPGPPPAEETTGRAKLOPPLPPE 720
 DB 661 LKSPWRGEYKPPHPPPNQANYHOTLHAQPRELSPRAPGPPPAEETTGRAKLOPPLPPE 720
 QY 721 PPEFNKSPPLTLKSESGMWEPLSLSELPAPARNPSSPERKATVPSEQLQQLBIELFLN 780
 DB 721 PPEFNKSPPLTLKSESGMWEPLSLSELPAPARNPSSPERKATVPSEQLQQLBIELFLN 780
 QY 781 SLQPPFSLERQEQILSLCLSDSLSDSEKNPSKASQSSRDITLSSGVHSSQAARSS 840
 DB 781 SLQPPFSLERQEQILSLCLSDSLSDSEKNPSKASQSSRDITLSSGVHSSQAARSS 840
 QY 841 SWNVLAGRPTDTPSYFNGVKVQIQLNGEHLHIREFHRVKGVDIATGISSQIPAAAFS 900
 DB 841 SWNVLAGRPTDTPSYFNGVKVQIQLNGEHLHIREFHRVKGVDIATGISSQIPAAAFS 900
 QY 901 LVTXGQGVRYDMEVPSGIDLOCTLAPDGSFAMSVKVGQLENRP 947
 DB 901 LVTXGQGVRYDMEVPSGIDLOCTLAPDGSFAMSVKVGQLENRP 947

RESULT 6

AAW82497
 ID AAW82497 standard; protein; 947 AA.

XX AC AAW82497;

XX DT 04-MAR-1999 (first entry)

XX DE Human NIK protein.

XX KW NIK; Nuclear factor-kappa B; NF-kB; signal transduction; TNF; human;
 KW tumour necrosis factor; NF-kB-inducing kinase; screening; interaction;
 KW modulator; immune response; inflammatory response; viral gene;

diagnostic; therapy.

KW
XX OS Homo sapiens.
XX PN US5854003-A.
XX XX 29-DEC-1998.
XX PD
XX XX
XX PF 26-FEB-1998; 98US-00032475.
XX XX 03-JUL-1997; 97US-00887518.
XX PR
XX PA (TULA-) TULARIK INC.
XX PI Wu L, Rothe M;
XX XX
XX DR WPI; 1999-094902/08.
XX DR N-PSDB; AAV73917.
XX XX
XX PT Screening agents for modulating interaction of nuclear factor kappaB
XX PT inducing kinase - with kinase-binding target, useful for controlling
XX PT levels of the kinase, for treatment and diagnosis of conditions
XX PT associated with e.g. inhibition of signal transduction by tumour necrosis
XX PT factor.
XX XX
XX PS Claim 6; Col 13-20; 16pp; English.
XX XX
XX CC This sequence represents a novel human nuclear factor-kappaB (NFkB)-
XX CC inducing kinase (NIK) which is used in a method for screening for agents
XX CC that modulate the interaction of NIK with a NIK-binding target. This
XX CC protein can be used as a modulator of cellular functions at the NIK
XX CC level, or for development of such compounds. NFkB is involved in
XX CC expression of many immune and inflammatory responses and of some
XX CC important viral genes. The protein may be used diagnostically and
XX CC therapeutically, in conditions associated with abnormal utilisation of
XX CC pathways that involve NFkB, e.g. inhibition of signal transduction by
XX CC tumour necrosis factor (TNF)
XX XX
XX SQ Sequence 947 AA;
Query Match 99.8%; Score 5044; DB 2; Length 947;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 946; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAVMEMACPGAGSAGVQKELPKKEKTPPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAGSAGVQKELPKKEKTPPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKEGSAGAAISIIAQAECECNSQBFSPFISRIIFAGSKQYSQESLDQIPNN 120
Db 61 VITKGTAKEGSAGAAISIIAQAECECNSQBFSPFISRIIFAGSKQYSQESLDQIPNN 120
QY 121 VAHATEGKMARVCKWKGKRSKARKKRSKSLAHAGVALAKPLPRTEQESCTIPVOE 180
Db 121 VAHATEGKMARVCKWKGKRSKARKKRSKSLAHAGVALAKPLPRTEQESCTIPVOE 180
QY 181 DESPLGAPYVNTPOFTKPLKEPGLQCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPYVNTPOFTKPLKEPGLQCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
QY 241 KLHPDQGGPLPLPHTPPYSLRPPHPPHPLQPKWHPLESFLGKLACVDSQKPLDDPH 300
Db 241 KLHPDQGGPLPLPHTPPYSLRPPHPPHPLQPKWHPLESFLGKLACVDSQKPLDDPH 300
QY 301 LSKLACVDSKPLPQPHLEPSCLSRGAEKFSVEEYLVHALQGSVSSQASHLSLAKTW 360
Db 301 LSKLACVDSKPLPQPHLEPSCLSRGAEKFSVEEYLVHALQGSVSSQASHLSLAKTW 360
QY 361 AARGSRSEPSKTEDNEGVLLTKLKPVDYVEEVEVHWATHQLRGSGFGEVHRMEDK 420
Db 361 AARGSRSEPSKTEDNEGVLLTKLKPVDYVEEVEVHWATHQLRGSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAELMACAGLTSPIVLYGAVREGPWNIIFMELLEGGSLGQL 480

Db 421 QTGFQCAVKVRLEVFRAELMACAGLTSPIVLYGAVREGPWNIIFMELLEGGSLGQL 480
QY 481 VKEOGCLPEDRALYYLGOALEGLEYLHSHRRILHGVKADNVLLSSDGGSHAALCDFGHAVC 540
Db 481 VKEOGCLPEDRALYYLGOALEGLEYLHSHRRILHGVKADNVLLSSDGGSHAALCDFGHAVC 540
QY 541 LOPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Db 541 LOPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
QY 601 RGPLCLKIASEPPPPVREIIPSCAPLTAQAIQEGLRKEPIHRSVAELGKGNRALQQVGG 660
Db 601 RGPLCLKIASEPPPPVREIIPSCAPLTAQAIQEGLRKEPIHRSVAELGKGNRALQQVGG 660
QY 661 LKSPWRGEYKEPRHPPPPNQANYHOTLHAQPRELSPRAFGPRPAEETTGRAFKLQPLPPE 720
Db 661 LKSPWRGEYKEPRHPPPPNQANYHOTLHAQPRELSPRAFGPRPAEETTGRAFKLQPLPPE 720
QY 721 PPEPNKSPPLTILSKESGMEPLDLSLEPAPARNPSSPERKATVPPEOLOOLIEIPLN 780
Db 721 PPEPNKSPPLTILSKESGMEPLDLSLEPAPARNPSSPERKATVPPEOLOOLIEIPLN 780
QY 781 SLSQPFSLSEOEQILSCLSIDSLSDSEKKNPKASQSSRDTLSGVSWSQAEARSS 840
Db 781 SLSQPFSLSEOEQILSCLSIDSLSDSEKKNPKASQSSRDTLSGVSWSQAEARSS 840
QY 841 SNNMVLARGRPTDTPSYFNGVKVQISLNGEHLHIREFHRVKVGDIAITGSSQIPAAAFS 900
Db 841 SNNMVLARGRPTDTPSYFNGVKVQISLNGEHLHIREFHRVKVGDIAITGSSQIPAAAFS 900
QY 901 LVTXGQGPVRYDMEVPDGSIDLOCTIAPDGSFAWSVRVKGQLENRP 947
Db 901 LVTXGQGPVRYDMEVPDGSIDLOCTIAPDGSFAWSVRVKGQLENRP 947
RESULT 7
AAW81564
ID AAW81564 standard; protein; 947 AA.
XX AC AAW81564;
XX DT 04-FEB-1999 (first entry)
XX DE Human NF-kB-inducing kinase (NIK) polypeptide.
XX KW NIK; IkB; NF-kB; TNF; kinase; nuclear factor kappa B; inhibition;
XX KW tumour necrosis factor; binding; genetic hybridisation; screening;
XX KW signal transduction; biopharmaceutical; human.
XX OS Homo sapiens.
XX XX US5843721-A.
XX PN 01-DEC-1998.
XX PD 03-JUL-1997; 97US-00887518.
XX PF 03-JUL-1997; 97US-00887518.
XX PR (TULA-) TULARIK INC.
XX PA Wu L, Rothe M;
XX PI
XX XX WPI; 1999-044580/04.
XX DR N-PSDB; AAV71603.
XX PT Probe, vector or recombinant nucleic acid encoding a polypeptide,
XX PT especially human nuclear factor kappa-B-inducing kinase protein - useful
XX PT for producing recombinant protein.
XX PS Claim 7; Col 13-18; 15pp; English.
XX XX

CC	This represents a nuclear factor kappa B (NF- κ B)-inducing kinase (NIK) polypeptide. The NIK polypeptide or its fragment has one or more activities selected from kinase activity and inhibitory activity; IkB kinase- α and beta binding activity and binding inhibitory activity; tumour necrosis factor (TNF) receptor-associated factor 2 binding activity and binding inhibitory activity; IkB binding activity and binding inhibitory activity, NF- κ B activating and inhibitory activity. A vector containing the NIK nucleic acid can be used to transform host cells for the recombinant production of the protein. The NIK nucleic acid and the polypeptide may be used in diagnosis (e.g. genetic hybridisation screen for NIK transcripts), therapy (e.g. NIK kinase inhibitors to inhibit TNF signal transduction), and in the biopharmaceutical industry									
XX										
SQ	Sequence 947 AA;									
Query Match		99.8%;		Score 5044;		DB 2;		Length 947;		
Best Local Similarity		99.9%;		Pred. No. 0;						
Matches 946;		Conservative 0;		Mismatches 1;		Indels 0;		Gaps 0;		
Qy	1	MAV	MEMACPGAGS	AVGQOKELPKPK	EKTPPLGKKQSSVYKLEAVEKSPVFCGKWEILND	60				
Db	1	MAV	MEMACPGAGS	AVGQOKELPKAK	EKTPPLGKKQSSVYKLEAVEKSPVFCGKWEILND	60				
Qy	61	VIT	KGTAKEGSEAG	PAISIIIAQAE	CENSQBFSTFSERIFIAGSKQYSQSESIDQIPNN	120				
Db	61	VIT	KGTAKEGSEAG	PAISIIIAQAE	CENSQBFSTFSERIFIAGSKQYSQSESIDQIPNN	120				
Qy	121	VAH	ATEGKMARVCW	KGRKRRKAR	KKKSKSLAHAGVALAKPLPRTPEQESCTIPVOE	180				
Db	121	VAH	ATEGKMARVCW	KGRKRRKAR	KKKSKSLAHAGVALAKPLPRTPEQESCTIPVOE	180				
Qy	181	DES	PLGAPYVRNT	PQFTKPLKE	PGLGQCFKQGLGELRPALPRSELHKLISPLQCLNHW	240				
Db	181	DES	PLGAPYVRNT	PQFTKPLKE	PGLGQCFKQGLGELRPALPRSELHKLISPLQCLNHW	240				
Qy	241	KLH	HPDGGELPL	PTHPPFYSRL	PHPPHPLQPKWPHPLESFLGKLACVDSQKPLDPDH	300				
Db	241	KLH	HPDGGELPL	PTHPPFYSRL	PHPPHPLQPKWPHPLESFLGKLACVDSQKPLDPDH	300				
Qy	301	LSK	LACVDSKPKPL	PGPHLEPCL	SRGAEKFSVEEYLVAHQGSVSSQAHSLTSLAKTW	360				
Db	301	LSK	LACVDSKPKPL	PGPHLEPCL	SRGAEKFSVEEYLVAHQGSVSSQAHSLTSLAKTW	360				
Qy	361	AAG	SRSRSPS	PKTDE	NEGVLITTEKLKPDVYREYEEVHWATHQLRGSGFGEVHRMEDK	420				
Db	361	AAG	SRSRSPS	PKTDE	NEGVLITTEKLKPDVYREYEEVHWATHQLRGSGFGEVHRMEDK	420				
Qy	421	QTG	QCQAVKVR	LVFVFAEEL	MACAGLTSPIVPLYGAVRGPWVNTFMELLEGSGLGOL	480				
Db	421	QTG	QCQAVKVR	LVFVFAEEL	MACAGLTSPIVPLYGAVRGPWVNTFMELLEGSGLGOL	480				
Qy	481	VKE	GCCLPEDRAL	YLYQAL	EGLEYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC	540				
Db	481	VKE	GCCLPEDRAL	YLYQAL	EGLEYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC	540				
Qy	541	LQP	DGLGKSL	LTGDI	IPGTETHMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF	600				
Db	541	LQP	DGLGKSL	LTGDI	IPGTETHMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF	600				
Qy	601	RGPL	CLKIASE	PPVPREI	PPSCAPLTAQAIQEGLRKEPIHRVSAEELGGKVNRA	660				
Db	601	RGPL	CLKIASE	PPVPREI	PPSCAPLTAQAIQEGLRKEPIHRVSAEELGGKVNRA	660				
Qy	661	LKS	PWRGEYKE	PRHP	PPNQANYHTLAQPRELSPRAGGPABETTCRAPKLOPLPPE	720				
Db	661	LKS	PWRGEYKE	PRHP	PPNQANYHTLAQPRELSPRAGGPABETTCRAPKLOPLPPE	720				
Qy	721	PPE	NKGPPL	TLT	SKESGMEPLPLSSLEPAPARNPSPERKATVPQELQQLIEFLN	780				
Db	721	PPE	NKGPPL	TLT	SKESGMEPLPLSSLEPAPARNPSPERKATVPQELQQLIEFLN	780				
Qy	781	SL	SQPF	SLSEGEQ	ILTCLSLSDSLSDSEKNPKSKASQSSRDITLSSGVHSSWSSQAEARSS	840				

Db	781	SLSQPFLERQEQILSLCLSLDSLSDDSEKNPKSKASQSSRDTLSSGVHSSWSSQEARSS	844			
Qy	841	SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRYKVGDIATGISQIPAAAFS	900			
Db	841	SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRYKVGDIATGISQIPAAAFS	900			
Qy	901	LVTKGQGPVRYDMEVPDSGIDLOCTLAPDGSFAWSRWYKHGQLENRP	947			
Db	901	LVTKGQGPVRYDMEVPDSGIDLOCTLAPDGSFAWSRWYKHGQLENRP	947			
RESULT 8						
AAW81561						
ID	AAW81561 standard; protein; 947 AA.					
XX	AAW81561;					
XX	04-FEB-1999 (first entry)					
DT	XX					
XX	XX					
DE	Human NF-kB-inducing kinase (NIK) polypeptide.					
XX	XX					
XX	NIK; IkB; NF-kB; TNF; kinase; nuclear factor kappa B; inhibition;					
KW	tumour necrosis factor; binding; genetic hybridisation; screening;					
KW	signal transduction; biopharmaceutical; immunogen; pharmacological;					
KW	transcription regulator; human.					
XX	XX					
OS	Homo sapiens.					
XX	XX					
PN	US5844073-A.					
XX	XX					
PD	01-DEC-1998.					
XX	XX					
PF	13-FEB-1998; 98US-00023321.					
XX	XX					
PR	03-JUL-1997; 97US-00887518.					
XX	XX					
PA	(TULA-) TULARIK INC.					
XX	XX					
PI	Wu L, Rothe M;					
XX	XX					
DR	WPI; 1999-044664/04.					
DR	N-PSDB; AAV69285.					
XX	XX					
PT	New isolated peptide comprising a specified 947 amino acid sequence - has					
PT	e.g. kinase activity, kinase inhibitory activity, IkB kinase-alpha					
PT	binding activity, and IkB kinase-alpha binding inhibitory activity.					
XX	XX					
PS	Claim 1; Col 13-18; 15pp; English.					
XX	XX					
CC	This represents a nuclear factor kappa B (NF-kB)-inducing kinase (NIK)					
CC	polypeptide. The NIK polypeptide or its fragment has one or more					
CC	activities selected from kinase activity and inhibitory activity; IkB					
CC	kinase-alpha and beta binding activity and binding inhibitory activity;					
CC	tumour necrosis factor (TNF) receptor-associated factor 2 binding					
CC	activity and binding inhibitory activity; IkB binding activity and					
CC	binding inhibitory activity, NF-kB activating and inhibitory activity.					
CC	The NIK nucleic acid and the polypeptide may be used in diagnosis (e.g.					
CC	genetic hybridisation screen for NIK transcripts), therapy (e.g. NIK					
CC	kinase inhibitors to inhibit TNF signal transduction), and in the					
CC	biopharmaceutical industry (e.g. as immunogens, reagents for isolating					
CC	other transcription regulators, and reagents for screening chemical					
CC	libraries for pharmacological agents)					
XX	XX					
SQ	Sequence 947 AA;					
Query Match 99.8%; Score 5044; DB 2; Length 947;						
Best Local Similarity 99.9%; Pred. No. 0;						
Matches 946; Conservative 0; Mismatches 1; Indels 0; Gaps 0;						
Qy	1	MAYMEMACPGAPGSAVQOQKELPPKPKETPLPGKKQSSVYKLEAVEKSPVFCGKWEILND	60			
Db	1	MAYMEMACPGAPGSAVQOQKELPPKPKETPLPGKKQSSVYKLEAVEKSPVFCGKWEILND	60			

```
QY 61 VITKGTAKESGAGPAAISIIAQAECENSQEFSTFSEIFIAAGSKQYSQESLDOI PNN 120
Db 61 VITKGTAKESGAGPAAISIIAQAECENSQEFSTFSEIFIAAGSKQYSQESLDOI PNN 120
QY 121 VAHATEGMARVCWKGKRSKARKKRSKSLAHAGVALAKPLRTPPEQSCCTIPVOE 180
Db 121 VAHATEGMARVCWKGKRSKARKKRSKSLAHAGVALAKPLRTPPEQSCCTIPVOE 180
QY 181 DESPLGAPYVRNTPOFTKPLKEPGLGQLCFKQEGELRPAIPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPYVRNTPOFTKPLKEPGLGQLCFKQEGELRPAIPRSELHKLISPLQCLNHVW 240
QY 241 KLHHPQDGGPLPLTHPPFPYSRLPHPPHPLQPKPPLPESFLGKLACVDSQKPLPDH 300
Db 241 KLHHPQDGGPLPLTHPPFPYSRLPHPPHPLQPKPPLPESFLGKLACVDSQKPLPDH 300
QY 301 LSKLACVDSQKPLPGPHLESCLSRGNAHEKFSVEEYLHVALQGSVSSQAHSLSLAKTW 360
Db 301 LSKLACVDSQKPLPGPHLESCLSRGNAHEKFSVEEYLHVALQGSVSSQAHSLSLAKTW 360
QY 361 AARGSRSPKPTDNEGVLTEKLPVDYREVEVHWATHQLRLGRSGFGEVHRMEDX 420
Db 361 AARGSRSPKPTDNEGVLTEKLPVDYREVEVHWATHQLRLGRSGFGEVHRMEDX 420
QY 421 QTGFQCAVKKVLEVFRAELMACAGLTSPRIPLYGAVREGPWNIFMELLEGSLGQL 480
Db 421 QTGFQCAVKKVLEVFRAELMACAGLTSPRIPLYGAVREGPWNIFMELLEGSLGQL 480
QY 481 VKEQGLCPEDRALYLGQALEGLEYLHSRIILHGDVKNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGLCPEDRALYLGQALEGLEYLHSRIILHGDVKNVLLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGKSLTGDIYIPGTETHMAPEVILGRSCDAKVDVWSSCCMWMLHMLNGCHPWTQFF 600
Db 541 LQPDGLGKSLTGDIYIPGTETHMAPEVILGRSCDAKVDVWSSCCMWMLHMLNGCHPWTQFF 600
QY 601 RGPLCLKIASPPVREIPSPCAPTAQAIQGLRKEPIHRSVAELGKGNRALQQVGG 660
Db 601 RGPLCLKIASPPVREIPSPCAPTAQAIQGLRKEPIHRSVAELGKGNRALQQVGG 660
QY 661 LKSPMRGEYKEPRHPPPNQANVHOTLHAQPRELSPRAPGPRPAETTTGRAPKLPPLPPE 720
Db 661 LKSPMRGEYKEPRHPPPNQANVHOTLHAQPRELSPRAPGPRPAETTTGRAPKLPPLPPE 720
QY 721 PEPNKSPLTLKSKEGWEPLPLSSLEPAPARNPSPERKATVPEQELQQLIEFLN 780
Db 721 PEPNKSPLTLKSKEGWEPLPLSSLEPAPARNPSPERKATVPEQELQQLIEFLN 780
QY 781 SLSQPFSLSEEQILSCLSIDSLSDSEKNPKASOSRDTLSSGVHSSQAERSS 840
Db 781 SLSQPFSLSEEQILSCLSIDSLSDSEKNPKASOSRDTLSSGVHSSQAERSS 840
QY 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGTSIQIPAAAFS 900
Db 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGTSIQIPAAAFS 900
QY 901 LVTKDGPVRVDMVDSGIDLQCTLAPDGSFAWSWRVKGQLENRP 947
Db 901 LVTKDGPVRVDMVDSGIDLQCTLAPDGSFAWSWRVKGQLENRP 947

RESULT 9
ID AAY31667
XX AAY31667 standard; protein; 947 AA.
AC AAY31667;
XX
XX
DT 09-NOV-1999 (first entry)
XX
DE NIK kinase-deficient mutant (T559A).
XX
KW NIK; NF-kB inducing kinase; inflammation; tumour necrosis factor;
interleukin-1; cytokine; inhibitor; antiinflammatory; apoptosis;
KW
```

```
KW anti-apoptotic; human; mutant.
XX
OS Homo sapiens.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 559 /note= "wild-type Thr is substituted by Ala"
FT
XX
XX WO9943704-Al.
XX
XX 02-SEP-1999.
XX
XX 25-FEB-1999; 99WO-US0041110.
XX
XX 27-FEB-1999; 98US-0076299P.
XX (REGC ) UNIV CALIFORNIA.
XX
XX Greene WC, Lin X, Gelezuinas R;
XX WPI; 1999-518837/43.
XX
XX New NF-kB inducing kinase proteins, used for inhibiting a kB-dependent
immune response, e.g. an inflammatory response or an anti-apoptotic
response.
XX
XX Claim 8; Page; 48pp; English.
XX
XX The present sequence represents a kinase-deficient mutant of human NF-kB
inducing kinase (NIK), in which the threonine residue at position 559 of
the wild-type protein (see AAY31665) is replaced by alanine. The
invention relates to: (a) an N-terminus deletion mutant NIK protein (see
AAY31666); and (b) a kinase deficient NIK mutant protein, especially the
present sequence of the sequence given in AAY31668, that inhibits auto-
phosphorylation or transphosphorylation. The invention provides the
molecular basis for cytokine induction of NF-kB-dependent immune and
inflammatory responses, emphasising a role for both NIK-NIK and NIK-IKK
(IkB-specific kinase) interactions. A novel and highly specific method
for modulating NF-kB-dependent immune, inflammatory and anti-apoptotic
responses is based on interruption of the critical protein interaction of
NIK and IKK. The mutant NIK proteins are used in claimed methods for
inhibiting NF-kB-dependent gene expression. The kinase-deficient NIK
mutant proteins inhibit the activation and phosphorylation of endogenous
IKK. Note: The present sequence is not shown in the specification but is
derived from the wild-type NIK sequence given in figure 10
XX
SQ Sequence 947 AA;
```

```
Query Match 99.8%; Score 5040; DB 2; Length 947;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAYWEMACGAPGAGSAGVQKQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAYWEMACGAPGAGSAGVQKQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKESGAGPAAISIIAQAECENSQEFSTFSEIFIAAGSKQYSQESLDOI PNN 120
Db 61 VITKGTAKESGAGPAAISIIAQAECENSQEFSTFSEIFIAAGSKQYSQESLDOI PNN 120
QY 121 VAHATEGMARVCWKGKRSKARKKRSKSLAHAGVALAKPLRTPPEQSCCTIPVOE 180
Db 121 VAHATEGMARVCWKGKRSKARKKRSKSLAHAGVALAKPLRTPPEQSCCTIPVOE 180
QY 181 DESPLGAPYVRNTPOFTKPLKEPGLGQLCFKQEGELRPAIPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPYVRNTPOFTKPLKEPGLGQLCFKQEGELRPAIPRSELHKLISPLQCLNHVW 240
QY 241 KLHHPQDGGPLPLTHPPFPYSRLPHPPHPLQPKPPLPESFLGKLACVDSQKPLPDH 300
Db 241 KLHHPQDGGPLPLTHPPFPYSRLPHPPHPLQPKPPLPESFLGKLACVDSQKPLPDH 300
```

Qy 301 LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVLHALQGSVSSQAHSLTSLAKTW 360
Db 301 LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVLHALQGSVSSQAHSLTSLAKTW 360
Qy 361 AARGSRSPKPTKEDNEGVLITLTKLPVDYREVEVHWATHQLRGRGSGFGEVHRMEDK 420
Db 361 AARGSRSPKPTKEDNEGVLITLTKLPVDYREVEVHWATHQLRGRGSGFGEVHRMEDK 420
Qy 421 QTGFQCAVKVRLVFRFAELMACAGLTSPRIPLVYGAVERGPWNIFMELLEGSGQL 480
Db 421 QTGFQCAVKVRLVFRFAELMACAGLTSPRIPLVYGAVERGPWNIFMELLEGSGQL 480
Qy 481 VKEQGCLPEDRALYILGQALEGLYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGCLPEDRALYILGQALEGLYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Qy 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Db 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Qy 601 RGPLCLKIASPPVREIPPSCAPLTAQAIOEGLRKEPIHRVSAELGGKVNRAALQQVGG 660
Db 601 RGPLCLKIASPPVREIPPSCAPLTAQAIOEGLRKEPIHRVSAELGGKVNRAALQQVGG 660
Qy 661 LKSPWRGEYKEPRHPPPNQANYHTLHAQPRELSPRAPGRPAEETTTGRAPKLPPLPPE 720
Db 661 LKSPWRGEYKEPRHPPPNQANYHTLHAQPRELSPRAPGRPAEETTTGRAPKLPPLPPE 720
Qy 721 PPEPNKSPPLTSLKEESGMWEPLSLLEPAPARNPSSPERKATVPQELQQLIEFLN 780
Db 721 PPEPNKSPPLTSLKEESGMWEPLSLLEPAPARNPSSPERKATVPQELQQLIEFLN 780
Qy 781 SLSQFFSLEEQEQLSCLSDSLSDSEKNPSKASQSSRDITLSSGVHSSWSSQAEARSS 840
Db 781 SLSQFFSLEEQEQLSCLSDSLSDSEKNPSKASQSSRDITLSSGVHSSWSSQAEARSS 840
Qy 841 SNNVLARGPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGSSQIPAAAFS 900
Db 841 SNNVLARGPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGSSQIPAAAFS 900
Qy 901 LVTKDGPVRYDMEVPSGIDLOCTLAPDGSFAWSRWKXGQLENRP 947
Db 901 LVTKDGPVRYDMEVPSGIDLOCTLAPDGSFAWSRWKXGQLENRP 947

RESULT 10
ABM85491
ID ABM85491 standard; protein; 947 AA.
AC ABM85491;
XX
XX
XX 18-NOV-2004 (first entry)
DT
XX
XX Human protein sequence hCP1763679.
DE
XX
XX Cytostatic; carcinoma; lymphoma; cancer; human.
XX
XX Homo sapiens.
OS
XX
XX W02003073826-A2.
XX
XX 12-SEP-2003.
XX
XX 28-FEB-2003; 2003WO-US0006235.
XX
XX 01-MAR-2002; 2002US-00087192.
FR
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW;
PI
XX
XX WPI; 2003-328604/31.
DR
XX

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX comprises a nucleotide sequence.
PS Claim 5; SEQ ID NO 888; Opp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CAP. Note:
CC This patent is an equivalent to basic patent US2002182586A1, for which no
CC sequence data was published
XX
SQ Sequence 947 AA;
Query Match 99.8%; Score 5040; DB 7; Length 947;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MAYMEMACPGAPGSAVQKQKELPKPEKTPPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAYMEMACPGAPGSAVQKQKELPKPEKTPPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Qy 61 VITKGTAKGSEAGPAAISIIAQAECNSQEFSPPTSERIFIAGSKQYSQSESIDQIPNN 120
Db 61 VITKGTAKGSEAGPAAISIIAQAECNSQEFSPPTSERIFIAGSKQYSQSESIDQIPNN 120
Qy 121 VAVATGEMARVCKWKGRSKARKKRSKSLAHAGVALAKPLPTPBOESCTIPVOE 180
Db 121 VAVATGEMARVCKWKGRSKARKKRSKSLAHAGVALAKPLPTPBOESCTIPVOE 180
Qy 181 DESPLGAPYVRNTPOFTKPLKEPGLQGLCFKQGLGELRPAIPSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPYVRNTPOFTKPLKEPGLQGLCFKQGLGELRPAIPSELHKLISPLQCLNHVW 240
Qy 241 KLHHPQDGGPLPLTPHPFPYSRLPHPPHPHPHPHPHPHPHPHPHPHPHPHPHPHP 300
Db 241 KLHHPQDGGPLPLTPHPFPYSRLPHPPHPHPHPHPHPHPHPHPHPHPHPHPHP 300
Qy 301 LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVLHALQGSVSSQAHSLTSLAKTW 360
Db 301 LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVLHALQGSVSSQAHSLTSLAKTW 360
Qy 361 AARGSRSPKPTKEDNEGVLITLTKLPVDYREVEVHWATHQLRGRGSGFGEVHRMEDK 420
Db 361 AARGSRSPKPTKEDNEGVLITLTKLPVDYREVEVHWATHQLRGRGSGFGEVHRMEDK 420
Qy 421 QTGFQCAVKVRLVFRFAELMACAGLTSPRIPLVYGAVERGPWNIFMELLEGSGQL 480
Db 421 QTGFQCAVKVRLVFRFAELMACAGLTSPRIPLVYGAVERGPWNIFMELLEGSGQL 480
Qy 481 VKEQGCLPEDRALYILGQALEGLYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGCLPEDRALYILGQALEGLYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Qy 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Db 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Qy 601 RGPLCLKIASPPVREIPPSCAPLTAQAIOEGLRKEPIHRVSAELGGKVNRAALQQVGG 660
Db 601 RGPLCLKIASPPVREIPPSCAPLTAQAIOEGLRKEPIHRVSAELGGKVNRAALQQVGG 660
Qy 661 LKSPWRGEYKEPRHPPPNQANYHTLHAQPRELSPRAPGRPAEETTTGRAPKLPPLPPE 720
Db 661 LKSPWRGEYKEPRHPPPNQANYHTLHAQPRELSPRAPGRPAEETTTGRAPKLPPLPPE 720

Qy 721 PPEPNKSPPLTLSEESGMWEPPLSLSELPAPARNPSSPERKATVPPEQLQLEIEFLN 780
 Db 721 PPEPNKSPPLTLSEESGMWEPPLSLSELPAPARNPSSPERKATVPPEQLQLEIEFLN 780
 Qy 781 SLSPQFSLERQEQILSLCLSDSLSDSEKPNFKASQSSRDTLSSGVHSSWSQAARSS 840
 Db 781 SLSPQFSLERQEQILSLCLSDSLSDSEKPNFKASQSSRDTLSSGVHSSWSQAARSS 840
 Qy 841 SNMWLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDITATGISSQIPAAAFS 900
 Db 841 SNMWLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDITATGISSQIPAAAFS 900
 Qy 901 LVTKDGPVRYDMEVDSGIDLOCTLAPDGSFAWSVRVXHGQLENRP 947
 Db 901 LVTKDGPVRYDMEVDSGIDLOCTLAPDGSFAWSVRVXHGQLENRP 947

RESULT 11
 AAY31668
 ID AAY31668 standard; protein; 947 AA.
 XX
 AC AAY31668;
 XX
 DT 09-NOV-1999 (first entry)
 XX
 DE NIK kinase-deficient mutant (S549A,T552A,T559A).
 XX
 KW NIK; NF-kB inducing kinase; inflammation; tumour necrosis factor;
 KW interleukin-1; cytokine; inhibitor; antiinflammatory; apoptosis;
 KW anti-apoptotic; human; mutant.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 549 /note= "wild-type Ser is substituted by Ala"
 FT Misc-difference 552 /note= "wild-type Thr is substituted by Ala"
 FT Misc-difference 559 /note= "wild-type Thr is substituted by Ala"
 XX
 XX WO943704-Al.
 XX
 XX 02-SEP-1999.
 XX
 XX 25-FEB-1999; 99WO-US004110.
 XX
 XX 27-FEB-1998; 98US-0076299P.
 XX
 XX (REGC) UNIV CALIFORNIA.
 XX
 XX Greene WC, Lin X, Gelezuinas R;
 XX WPI; 1999-518837/43.
 XX
 XX New NF-kB inducing kinase proteins, used for inhibiting a kB-dependent
 XX immune response, e.g. an inflammatory response or an anti-apoptotic
 XX response.
 XX
 XX Claim 9; Page; 48pp; English.
 XX
 XX The present sequence represents a kinase-deficient mutant of human NF-kB
 XX inducing kinase (NIK), in which the serine residue at position 549 of the
 XX wild-type protein (see AAY31665) is replaced by alanine, the threonine
 XX residue at position 552 is replaced by alanine, and the threonine residue
 XX at position 559 is replaced by alanine. The invention relates to: (a) an
 XX N-terminus deletion mutant NIK protein. (see AAY31666); and (b) a kinase
 XX deficient NIK mutant protein, especially the present sequence of the
 XX sequence given in AAY31667, that inhibits auto-phosphorylation or
 XX transphosphorylation. The invention provides the molecular basis for
 XX cytokine induction of NF-kB-dependent immune and inflammatory responses,

CC emphasising a role for both NIK-NIK and NIK-IKK (IkB-specific kinase)
 CC interactions. A novel and highly specific method for modulating NF-kB-
 CC dependent immune, inflammatory and anti-apoptotic responses is based on
 CC interruption of the critical protein interaction of NIK and IKK. The
 CC mutant NIK proteins are used in claimed methods for inhibiting NF-kB-
 CC dependent gene expression. The kinase-deficient NIK mutant proteins
 CC inhibit the activation and phosphorylation of endogenous IKK. Note: The
 CC present sequence is not shown in the specification but is derived from
 CC the wild-type NIK sequence given in figure 10
 XX
 SQ Sequence 947 AA;
 Query Match 99.6%; Score 5032; DB 2; Length 947;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 943; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 MAVMEMACPGAGPGSAVGQKELPKPKETPLPKKQSSVYKLEAVEKSPVFCGKWEILND 60
 Db 1 MAVMEMACPGAGPGSAVGQKELPKPKETPLPKKQSSVYKLEAVEKSPVFCGKWEILND 60
 Qy 61 VITKTAKEGSEAGPAAISIIAQACENSOQSFSTPFSERIFIAGSKOYSQSESIDQIPNN 120
 Db 61 VITKTAKEGSEAGPAAISIIAQACENSOQSFSTPFSERIFIAGSKOYSQSESIDQIPNN 120
 Qy 121 VAHATEGMARVCWKGRSKARKKRRKSKSLAHAGVALAKPLPRTPEQESCITIPVOE 180
 Db 121 VAHATEGMARVCWKGRSKARKKRRKSKSLAHAGVALAKPLPRTPEQESCITIPVOE 180
 Qy 181 DESPLGAPYVRNTPQFTKPLKEPGLQCFKQLEGRLPALPRSELHKLSPLOCLNHVW 240
 Db 181 DESPLGAPYVRNTPQFTKPLKEPGLQCFKQLEGRLPALPRSELHKLSPLOCLNHVW 240
 Qy 241 KLHPDQGGPLPLPHTPFPYSRLPHPPHPLQPKWPHPLESFLGKLACVDSQKPLDPH 300
 Db 241 KLHPDQGGPLPLPHTPFPYSRLPHPPHPLQPKWPHPLESFLGKLACVDSQKPLDPH 300
 Qy 301 LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQASHLTSIAKTW 360
 Db 301 LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQASHLTSIAKTW 360
 Qy 361 AARGSRSPSPKTEDNEGVLLETKLPVDYREEVHMAHQLRLGRGSGFGEVHRMEDK 420
 Db 361 AARGSRSPSPKTEDNEGVLLETKLPVDYREEVHMAHQLRLGRGSGFGEVHRMEDK 420
 Qy 421 QTGFQCAVKVRLEVFRAELMACAGLTSPRIYPLGAVREGFWNIFMELLEGGSLGQL 480
 Db 421 QTGFQCAVKVRLEVFRAELMACAGLTSPRIYPLGAVREGFWNIFMELLEGGSLGQL 480
 Qy 481 VKEQGCLPEDRALYILGQALEGLEYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
 Db 481 VKEQGCLPEDRALYILGQALEGLEYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
 Qy 541 LQPDGLGKSLITGDYIPGTETHMADEVVLGRSCDAKVDVWSSCCMLHMLNGCHPWTQFF 600
 Db 541 LQPDGLGKSLITGDYIPGTETHMADEVVLGRSCDAKVDVWSSCCMLHMLNGCHPWTQFF 600
 Qy 601 RGPLCLKIASPPVREIPPSCAPLTAQAOIEGLRKEPIHRVSAAEELGKVNALQOVGG 660
 Db 601 RGPLCLKIASPPVREIPPSCAPLTAQAOIEGLRKEPIHRVSAAEELGKVNALQOVGG 660
 Qy 661 LKSPWRGEYKEPRHPPPNQANYHTLHAQPRELSPRAPGPRPAEETTGAPKLQPLPPE 720
 Db 661 LKSPWRGEYKEPRHPPPNQANYHTLHAQPRELSPRAPGPRPAEETTGAPKLQPLPPE 720
 Qy 721 PPEPNKSPPLTLSEESGMWEPPLSLSELPAPARNPSSPERKATVPPEQLQLEIEFLN 780
 Db 721 PPEPNKSPPLTLSEESGMWEPPLSLSELPAPARNPSSPERKATVPPEQLQLEIEFLN 780
 Qy 781 SLSPQFSLERQEQILSLCLSDSLSDSEKPNFKASQSSRDTLSSGVHSSWSQAARSS 840
 Db 781 SLSPQFSLERQEQILSLCLSDSLSDSEKPNFKASQSSRDTLSSGVHSSWSQAARSS 840
 Qy 841 SNMWLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDITATGISSQIPAAAFS 900

Db 841 SNNWLAGRPTDPSYNGVKVQSQUNGHLEHIREFHRVVKVGDIAIGSSQIPAAAFS 900
Qy 901 LVTKDQGVRYDMEVPDSDGLDQCTLAPDGSFANRWVKHGQLENRP 947
Db 901 LVTKDQGVRYDMEVPDSDGLDQCTLAPDGSFANRWVKHGQLENRP 947
RESULT 12
ADP29813
ID ADP29813 standard; protein; 697 AA.
XX AC ADP29813;
XX DT 12-AUG-2004 (first entry)
XX DE Human secreted protein SEQ ID #580.
XX KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; Inflammatory; Immune; human secreted protein.
XX OS Homo sapiens.
XX PN WO2004035732-A2.
XX PD 29-APR-2004.
XX PF 28-AUG-2003; 2003WO-US026780.
XX PP 29-AUG-2002; 2002US-0406576P.
XX PR 29-AUG-2002; 2002US-0406579P.
XX PR 29-AUG-2002; 2002US-0406585P.
XX PR 29-AUG-2002; 2002US-0406588P.
XX PR 29-AUG-2002; 2002US-0406608P.
XX PR 29-AUG-2002; 2002US-0406611P.
XX PR 29-AUG-2002; 2002US-0406612P.
XX PR 29-AUG-2002; 2002US-0406616P.
XX PR 29-AUG-2002; 2002US-0406640P.
XX PR 29-AUG-2002; 2002US-0406642P.
XX PR 29-AUG-2002; 2002US-0406646P.
XX PR 29-AUG-2002; 2002US-0406653P.
XX PR 29-AUG-2002; 2002US-0406655P.
XX PR 29-AUG-2002; 2002US-0406666P.
XX PR 17-SEP-2002; 2002US-0410946P.
XX PR 17-SEP-2002; 2002US-0410947P.
XX PR 17-SEP-2002; 2002US-0410948P.
XX PR 17-SEP-2002; 2002US-0410949P.
XX PR 17-SEP-2002; 2002US-0410953P.
XX PR 17-SEP-2002; 2002US-0410957P.
XX PR 17-SEP-2002; 2002US-0410958P.
XX PR 17-SEP-2002; 2002US-0410959P.
XX PR 17-SEP-2002; 2002US-0410960P.
XX PR 17-SEP-2002; 2002US-0410961P.
XX PR 17-SEP-2002; 2002US-0410962P.
XX PR 17-SEP-2002; 2002US-0411019P.
XX PR 17-SEP-2002; 2002US-0411022P.
XX PR 17-SEP-2002; 2002US-0411023P.
XX PR 17-SEP-2002; 2002US-0411024P.
XX PR 17-SEP-2002; 2002US-0411032P.
XX PR 17-SEP-2002; 2002US-0411035P.
XX PR 17-SEP-2002; 2002US-0411037P.
XX PR 17-SEP-2002; 2002US-0411041P.
XX PR 17-SEP-2002; 2002US-0411045P.
XX PR 17-SEP-2002; 2002US-0411046P.
XX PR 17-SEP-2002; 2002US-0411048P.
XX PR 17-SEP-2002; 2002US-0411052P.
XX PR 17-SEP-2002; 2002US-0411055P.
XX PR 17-SEP-2002; 2002US-0411073P.
XX PR 17-SEP-2002; 2002US-0411082P.
XX PR 17-SEP-2002; 2002US-041101P.
XX PR 17-SEP-2002; 2002US-041111P.
XX PR 18-APR-2003; 2003US-0463700P.
XX PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467210P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 14-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
XX PI Halenbeck RP, Huang WM, Kothakota S, Haishan L, Linnemann T;
XX PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.
XX PS Claim 1; SEQ ID NO 1811; 428pp; English.
XX PS The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytostatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPWEB and is not in the specification.
SQ Sequence 697 AA;
Query Match 68.1%; Score 3439.5; DB 8; Length 697;
Best Local Similarity 72.7%; Pred. No. 5.8e-246;
Matches 688; Conservative 0; Mismatches 2; Indels 257; Gaps 5;
Qy 1 MAVNEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGRWEILND 60
Db 1 MAVNEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGRWEILND 60
Qy 61 VITKGTAKGSEAGPAAISIIAQACENSQEFSTFSERIFIAGSKQYSQSESLDQIPNN 120
Db 61 VITKGTAKGSEAGPAAISIIAQACENSQEFSTFSERIFIAGSKQYSQSESLDQIPNN 120
Qy 121 VAHATEGMARVCKWKGKRRSKARKKRSKSLAHAGVALAKPLPTPEQESTCTIPVQ 180
Db 121 VAHATEGMARVCKWKGKRRSKARKKRSKSLAHAGVALAKPLPTPEQESTCTIPVQ 179
Qy 181 DESPLGAPYVRNTPOFTKPLKEPLGQLCFKQLGEGRLPALPRSELHKLISPLQLNHW 240
Db 180 ----- 179
Qy 241 KLHHPDQGGPLPLPTHPPFYSRPLPHPPFHPQWKPHPLESFLGKLCACVDSQKPLDPH 300
Db 180 ----- 179

QY 301 LSKLACVDSKPLPGPHLEPCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLTSLAKTW 360
 Db 180 -----LPGPHLEPCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLTSLAKTW 227
 QY 361 AARGSRSPSPKPTDNEGVLTEKLKPDYDYEYREEVHWATHQLRLGRSGFGEVHRMEDK 420
 Db 228 AARGSRSPSPKPTDNEGVLTEKLKPDYDYEYREEVHWATHQLRLGRSGFGEVHRMEDK 287
 QY 421 QTGFQCAVKKVRLEVFRAELMACAGLTSPRIPLVYCAVREGPWNIIFMELLEGGSLGOL 480
 Db 288 QTGFQCAVKK-----YCAVREGPWNIIFMELLEGGSLGOL 322
 QY 481 VKEQGLCPEDRALYILGQALEGLEYLHSRRLHGDVKADNVLSSDGSAAALCDFGHAVC 540
 Db 323 VKEQGLCPEDRALYILGQALEGLEYLHSRRLHGDVKADNVLSSDGSAAALCDFGHAVC 382
 QY 541 LPDGLGKSLLTGDIYIPGTETHMAPEVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600
 Db 383 LPDGLGKSLLTGDIYIPGTETHMAPEVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 442
 QY 601 RGPLCLKIASPPVREIIPPSCAPLTAQAIQGLRKEPIHRVSAABELGCKVNRALQOVGG 660
 Db 443 RGPLCLKIASPPVREIIPPSCAPLTAQAIQGLRKEPIHRVSAABELGCKVNRALQOVGG 499
 QY 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSRAPGPRPAEETTGRAPKLQPPLPPE 720
 Db 500 -----E 500
 QY 721 PPEPNKSPPLTLKSKEGSMWEPLPLSSLEPAPARNPSPERKATVPEQELQLELFLN 780
 Db 501 PPEPNKSPPLTLKSKEGSMWEPLPLSSLEPAPARNPSPERKATVPEQELQLEI---LN 557
 QY 781 SLSQPFSLEEQIILSCLSIDSLSDSEKNPKSASQSSRDTLSSGVHSSQAEARSS 840
 Db 558 SLSQPFSLEEQIILSCLSIDSLSDSEK----- 588
 QY 841 SNNMVLARGRPDTTFSYFNGVKVQIQSLNGEHLHIREFHRVKGVDIATGISQIPAAAFS 900
 Db 589 -----LARGRPDTTFSYFNGVKVQIQSLNGEHLHIREFHRVKGVDIATGISQIPAAAFS 643
 QY 901 LVTGQGPVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKHGQLENRP 947
 Db 644 LVTGQGPVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKHGQLENRP 690

RESULT 13
 ABM85490
 ID ABM85490 standard; protein; 548 AA.
 AC ABM85490;
 XX 18-NOV-2004 (first entry)
 DT Mouse protein sequence mCP14606.
 DE Cytostatic; carcinoma; lymphoma; cancer; murine.
 KW Mus musculus.
 OS Mus musculus.
 XX WO2003073826-A2.
 FN 28-FEB-2003; 2003WO-US006235.
 XX 12-SEP-2003.
 PD 28-FEB-2003; 2003WO-US006235.
 XX 01-MAR-2002; 2002US-00087192.
 PR (SAGR-) SAGRES DISCOVERY.
 XX Morris DW;
 XX WPI; 2003-328604/31.

XX PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
 PT comprises a nucleotide sequence.
 XX Claim 5; SEQ ID NO 885; Opp; English.
 PS The present invention relates to novel DNA and protein sequences which
 XX are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
 CC determining Carcinoma Associated (CA) gene copy number. In addition, the
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
 CC carcinoma including lymphoma. The present sequence is one such CAP. Note:
 CC This patent is an equivalent to basic patent US2002182586A1, for which no
 CC sequence data was published
 XX Sequence 548 AA;
 SQ

Query Match 49.1%; Score 2481.5; DB 7; Length 548;
 Best Local Similarity 86.1%; Pred. No. 4.9e-175;
 Matches 466; Conservative 32; Mismatches 36; Indels 7; Gaps 2;

QY 407 GRGSFGEVHRMEDKOTGFQCAVKKVRLEVFRAELMACAGLTSPRIPLVYCAVREGPWNI 466
 Db 15 QSSSFGFVHRMKDQTFQCAVKKVRLEVFRAELMACAGLTSPRIPLVYCAVREGPWNI 74
 QY 467 IFMELLEGGSLGOLVKQGLCPEDRALYILGQALEGLEYLHSRRLHGDVKADNVLSSD 526
 Db 75 IFMELLEGGSLGOLVKQGLCPEDRALYILGQALEGLEYLHSRRLHGDVKADNVLSSD 134
 QY 527 GSHAALCDFGHAVCLQPDGLGKSLLTGDIYIPGTETHMAPEVLGRSCDAKVDVWSSCCMW 586
 Db 135 GSRAALCDFGHAVCLQPDGLGKSLLTGDIYIPGTETHMAPEVLGRSCDAKVDVWSSCCMW 194
 QY 587 LHMNGCHPWTQFPRGPLCKIASPPVREIIPPSCAPLTAQAIQGLRKEPIHRVSAAE 646
 Db 195 LHMNGCHPWTQFPRGPLCKIASPPVREIIPPSCAPLTAQAIQGLRKEPIHRVSAAE 254
 QY 647 LGGKVNALQOVGKLSKSPWGEYKEPRHPPPNQANYHOTLHAQPRELSRAPGPRPAEET 706
 Db 255 LRRKVGKALQOVGKLSKSPWGEYKEPRHPPPNQANYHOTLHAQPRELSRAPGPRPAEET 307
 QY 707 TGRAPKLQPPLEPPPEPNKSPPLTLKSKEGSMWEPLPLSSLEPAPARNPSPERKATVP 766
 Db 308 TDGAPEQPPLPPEPPPEPNKSPALNLSKESGTWEPPLSSLDPATAGSFDPDRATLP 367
 QY 767 EQELOQLEIEFLNLSLQSPFSLBEQOILSCLSIDSLSDSEKNPKSASQSSRDTLSS 826
 Db 368 EQELOQLEIEFLNLSLQSPFSLBEQOILSCLSIDSLSDSEKNPKSASQSSRDTLSS 427
 QY 827 GVHSSWSQAEARSSNNMVLARGRPDTTFSYFNGVKVQIQSLNGEHLHIREFHRVKGDI 886
 Db 428 GVHSSWSQAEARSSNNMVLARGRPDTTFSYFNGVKVQIQSLNGEHLHIREFHRVKGDI 487
 QY 887 ATGISSQIPAAAFSLVTGQGPVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKHGQLENR 946
 Db 488 ATGISSQIPATFSLVTGQGPVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKHGQLENR 547
 QY 947 P 947
 Db 548 P 548

RESULT 14
 ADQ67717
 ID ADQ67717 standard; protein; 477 AA.
 XX ADQ67717;
 AC

XX 07-OCT-2004 (first entry)
XX Novel human protein sequence #2383.
XX osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX Homo sapiens.
XX EPI440981-A2.
XX 28-JUL-2004.
XX 21-JAN-2004; 2004EP-00001196.
XX 21-JAN-2003; 2003JP-00102206.
PR 09-MAY-2003; 2003JP-00131392.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
XX WPI; 2004-535376/52.
DR N-PSDB; ADQ67410.
XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX Claim 1; SEQ ID NO 4878; 2449pp; English.
XX The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a protein
CC sequence of the invention.
XX SQ Sequence 477 AA;
Query Match 45.5%; Score 2301; DB 8; Length 477;
Best Local Similarity 99.3%; Pred. No. 9.8e-162;
Matches 431; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 514 GDVKNVLLSSDGHAAALCDPFGHACVCLQPPGLGKSLTGDYIPCTETHMAPEVVLGRSC 573
DB : |||||
44 GGLAANDVLLSSDGHAAALCDPFGHACVCLQPPGLGKSLTGDYIPCTETHMAPEVVLGRSC 103
QY |||||
574 DAKVDVSSCCMLHMLNGCHPWTQFFRGPLCLKIASEPPPPVREIPSPCAPLTAQIQEG 633
DB |||||
104 DAKVDVSSCCMLHMLNGCHPWTQFFRGPLCLKIASEPPPPVREIPSPCAPLTAQIQEG 163
QY |||||
634 LRKEPIHRVSAELGGKVNALQOVGGLKSPWRGEYKEPRHPPPNQANYHOTLHAQPREL 693
DB |||||
164 LRKEPIHRVSAELGGKVNALQOVGGLKSPWRGEYKEPRHPPPNQANYHOTLHAQPREL 223
QY |||||
694 SPRAGPRPAETTCRAPKLOPPLPPEPPENKSPPLTLSEESGMWEPLLSLEPAPA 753
DB |||||
224 SPRAGPRPAETTCRAPKLOPPLPPEPPENKSPPLTLSEESGMWEPLLSLEPAPA 283
QY |||||
754 RNPSSPERKATVPQEQLOQLIEFLNLSLSPFSLSEEQILSLCLSDLSLSDSEKNP 813
DB |||||
284 RNPSSPERKATVPQEQLOQLIEFLNLSLSPFSLSEEQILSLCLSDLSLSDSEKNP 343
QY |||||
814 SKAQSSRDITLSSGVHSSQAEARSSSNMVLARGRPTDTPSYFNGVKVQIQSLNGEHL 873
DB |||||
344 SKAQSSRDITLSSGVHSSQAEARSSSNMVLARGRPTDTPSYFNGVKVQIQSLNGEHL 403

QY 874 HIREFHRVKVGDIAITGISSQIPAAAFSLVTXGQPVRYDMEVPDSDIGDLQCTLAPDGSFA 933
DB |||||
404 HIREFHRVKVGDIAITGISSQIPAAAFSLVTXGQPVRYDMEVPDSDIGDLQCTLAPDGSFA 463
QY 934 WSWRVKVGQLENRP 947
DB |||||
464 WSWRVKVGQLENRP 477
RESULT 15
ABR61616
ID ABR61616 standard; protein; 324 AA.
XX AC ABR61616;
XX 15-JAN-2004 (first entry)
XX Human NIK C-terminal fragment (residues 624-947).
XX NIK; IL-2; interleukin 2; common gamma chain; cgammac; antiinflammatory;
KW gastrointestinal; antiarthritic; antirheumatic; osteopathic; cardiant;
KW antiasthmatic; nootropic; neuroprotective; antiarteriosclerotic; NIKK;
KW immunosuppressive; antithyroid; nuclear factor KB inducing kinase;
KW gene therapy; antianemic; NIK kinase; human.
XX Homo sapiens.
XX WO2003087380-A1.
XX 23-OCT-2003.
XX 15-APR-2003; 2003WO-IL000317.
XX 18-APR-2002; 2002IL-00149217.
PR 08-OCT-2002; 2002IL-00152183.
XX (YEDA) YEDA RES & DEV CO LTD.
XX Wallach D, Ramakrishnan P, Shmushkovich T;
XX WPI; 2003-845333/78.
XX New nuclear factor inducing kinase or its mutein, variant, fusion
PT protein, functional derivative, circularly permuted derivative or
PT fragment, useful for treating an autoimmune disease, infarct, Alzheimer's
PT disease or atherosclerosis.
XX Claim 2; Page 92-93; Opp; English.
XX The invention relates to a NIK (nuclear factor (NF)-KB-inducing kinase)
CC or its mutein, variant, fusion protein, functional derivative, circularly
CC permuted derivative or fragment. Specific antibodies and small
CC molecules capable of modulating the interaction between interleukin 2 (IL
CC -2) common gamma chain (cgammac) and NIK kinase (NIKK) are useful for the
CC manufacture of a medicament for the treatment of a disease, e.g. a
CC disease resulting from excessive immune response such as rheumatoid
CC arthritis, osteoarthritis, inflammatory bowel disease, asthma, cardiac
CC infarct, Alzheimer's disease or atherosclerosis; or an autoimmune disease
CC such as immune thyroiditis, or other arthropathies, such as autoimmune
CC haemolytic anemia. The small molecule is useful for modulating signaling
CC through cgammac. The present sequence represents a human NIK fragment
XX SQ Sequence 324 AA;
Query Match 33.7%; Score 1703; DB 7; Length 324;
Best Local Similarity 100.0%; Pred. No. 1.3e-117;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 624 PLTAAQIQEGLRKEPIHRVSAELGGKVNALQOVGGLKSPWRGEYKEPRHPPPNQANYH 683
DB |||||
1 PLTAAQIQEGLRKEPIHRVSAELGGKVNALQOVGGLKSPWRGEYKEPRHPPPNQANYH 60

QY 684 QTLHAQPRELSPRACGPRPAEETTCRAPKLQPPPLPEPPEPNKSPPLTLKSKESGMWBPL 743
Db |||||
61 QTLHAQPRELSPRACGPRPAEETTCRAPKLQPPPLPEPPEPNKSPPLTLKSKESGMWBPL 120
QY 744 PLSSLEPAPARNPSSPERKATVPEQELQOLEIELEFLNLSQPFSLBEOHILSCLSIDSL 803
Db |||||
121 PLSSLEPAPARNPSSPERKATVPEQELQOLEIELEFLNLSQPFSLBEOHILSCLSIDSL 180
QY 804 SLSDDEKNPKSKASOSSRDTLSSGVHWSQAEARSSWNMYLARGRPTDTPSYFNGVKV 863
Db |||||
181 SLSDDEKNPKSKASOSSRDTLSSGVHWSQAEARSSWNMYLARGRPTDTPSYFNGVKV 240
QY 864 QIQSLNGEHLHIREFHRVKVGDIAIGISSQIPAAAFSLVTKDQGPVRYDMEVPDGSIDIQ 923
Db |||||
241 QIQSLNGEHLHIREFHRVKVGDIAIGISSQIPAAAFSLVTKDQGPVRYDMEVPDGSIDIQ 300
QY 924 CTLAPDGSFAMSWRVKHGQLENRP 947
Db |||||
301 CTLAPDGSFAMSWRVKHGQLENRP 324

Search completed: June 7, 2005, 11:57:28
Job time : 150.36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:40:40 ; Search time 138.585 Seconds
(without alignments)
3499.208 Million cell updates/sec

Title: US-09-155-676B-7

Perfect score: 5052

Sequence: 1 MAVNEMACPGAGSAGCQK.....PDGSFAWSRVKHCQLENRP 947

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03:*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5052	100.0	947	1 M3KE_HUMAN	Q9558 homo sapien
2	5040	99.8	947	2 Q1Y1N1	Q81yn1 homo sapien
3	4196.5	83.1	942	1 M3KE_MOUSE	Q9wul6 mus musculu
4	2301	45.5	477	2 Q6ZM21	Q6zmz1 homo sapien
5	2110	41.8	433	2 Q68D39	Q88d39 homo sapien
6	360	7.1	467	1 M3K8_HUMAN	P41279 homo sapien
7	360	7.1	467	2 Q6FG25	Q6fg25 homo sapien
8	357.5	7.1	467	1 M3K8_MOUSE	Q07174 mus musculu
9	356.5	7.1	467	1 M3K8_RAT	Q63562 rattus norv
10	356	7.0	1291	2 Q9WT82	Q9wtr2 mus musculu
11	355.5	7.0	1493	1 M3K1_MOUSE	P33349 mus musculu
12	353	7.0	619	1 M3K2_MOUSE	Q61083 mus musculu
13	353	7.0	1493	1 M3K1_RAT	Q62925 rattus norv
14	351	6.9	1158	2 Q6C2G1	Q6c2g1 yarrowia li
15	350.5	6.9	619	2 Q9NYK3	Q9nyk3 homo sapien
16	346	6.8	423	2 Q9NC32	Q9nc32 homo sapien
17	346	6.8	423	2 Q9NC32	Q9nc32 homo sapien
18	345	6.8	618	1 M3K1_HUMAN	Q32333 homo sapien
19	343.5	6.8	614	2 Q6RFY3	Q9y2u5 homo sapien
20	341.5	6.8	944	2 Q6C1T1	Q6rfy3 lycopersico
21	341.5	6.8	944	2 Q6C1T1	Q6clt1 yarrowia li
22	341	6.7	1320	2 Q9UCV9	Q7z8j5 yarrowia li
23	338.5	6.7	611	2 Q6RFY4	Q9ucv9 emericella
24	335.5	6.6	690	2 Q8H334	Q8rfy4 nicotiana b
25	332.5	6.6	894	2 Q7XU11	Q8h334 oryza sativ
26	332	6.6	735	2 Q6Z189	Q7xur1 oryza sativ
27	327	6.5	709	2 Q7XV58	Q6zi89 oryza sativ
28	326.5	6.5	651	1 M3K3_ARATH	Q7xv58 oryza sativ
29	325	6.4	651	1 M3K2_ARATH	O22042 arabidopsis
30	325	6.4	1288	2 Q75PK3	Q9fz36 arabidopsis
31	325	6.4	1478	1 BCK1_YEAST	Q01389 saccharomyc

RESULT 1

ID M3KE_HUMAN STANDARD; PRT; 947 AA.

AC Q9558;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 25-JAN-2005 (Rel. 46, Last annotation update)

DE Mitogen-activated protein kinase kinase 14 (EC 2.7.1.37) (NP-kappa beta-inducing kinase) (Serine/threonine-protein kinase NIK) (HaNIK)

DE (Name=MAP3K14; Synonyms=NIK;

GN Homo sapiens (Human)

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND MUTAGENESIS OF 429-LYS-LYS-430.

RX MEDLINE=97172277; PubMed=9020361;

RA Malinin N.L., Boldin M.P., Kovalenko A.V., Wallach D.;

RT "MAP3K-related kinase involved in NF-kappaB induction by TNF, CD95 and IL-1.";

RL Nature 385:540-544 (1997).

RN [2]

RP INTERACTION WITH PELI3.

RX MEDLINE=22756745; PubMed=12874243;

RA Jensen L.E., Whitehead A.S.;

RT "Pellino3, a novel member of the Pellino protein family, promotes activation of c-Jun and Elk-1 and may act as a scaffolding protein.";

RL J. Immunol. 171:1500-1506(2003).

CC -!- FUNCTION: Lymphotoxin beta-activated kinase which seems to be exclusively involved in the activation of NF-kappa-B and its transcriptional activity. Induces the processing of NF-kappa-B 2/P100. Could act in a receptor-selective manner (By similarity).

CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.

CC -!- SUBUNIT: Binds to TRAF2, TRAF5, TRAF6, IKKA and NF-kappa-B 2/P100 (By similarity). Interacts with PELI3.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- TISSUE SPECIFICITY: Weakly expressed in testis, small intestine, spleen, thymus, peripheral blood leukocytes, prostate, ovary and colon.

CC -!- PTM: Autophosphorylated.

CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP kinase kinase kinase subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).

CC or send an email to license@isb-sib.ch.

CC -----

CC HESB; Y10256; CAA71306.1; -.

CC HSSP; P71584; 1MRU.

DR

ALIGNMENTS

32 324.5 6.4 1497 2 Q95YH6 Q95yh6 drosophila

33 324.5 6.4 1571 2 Q8MSQ4 Q8meq4 drosophila

34 324.5 6.4 1571 2 Q95YH7 Q95yh7 drosophila

35 324.5 6.4 1612 2 Q9VE37 Q9ve37 drosophila

36 323 6.4 797 2 Q9EB75 Q9eb75 homo sapien

37 323 6.4 1011 1 M3K6_HUMAN Q95382 homo sapien

38 322.5 6.4 894 2 Q6ESH1 Q6esh1 oryza sativ

39 322 6.4 1311 2 Q9ULE2 Q9ule2 homo sapien

40 322 6.4 1315 2 Q8TC32 Q8tc32 homo sapien

41 322 6.4 1315 2 Q9NRP7 Q9nrp7 homo sapien

42 321.5 6.4 674 2 Q6H602 Q6h602 oryza sativ

43 320.5 6.3 650 2 Q95ZN5 Q95zn5 caenorhabdi

44 320 6.3 839 2 Q8NUX3 Q8njx3 ustilago ma

45 319.5 6.3 591 2 Q82667 Q82667 brassica na

DR IntAct; Q99558; --
DR Genew; HGNC:6853; MAP3K14.
DR MTM; 604655; --
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Phosphorylation; Serine/threonine-protein kinase;
KW Transferase.
FT DOMAIN 400 655 Protein kinase.
FT NP_BIND 406 414 ATP (By similarity).
FT BINDING 429 429 ATP (By similarity).
FT ACT_SITE 515 515 Proton acceptor (By similarity).
FT MUTAGEN 429 430 KK->AA: Loss of autophosphorylation.
SQ SEQUENCE 947 AA; 104097 MW; D307F76993BC9661 CRC64;
Query Match 100.0%; Score 5052; DB 1; Length 947;
Best Local Similarity 100.0%; Pred. No. 6.5e-241;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAVMEMACPGAGSAGVQOKELPKPKETPLPKGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAGSAGVQOKELPKPKETPLPKGKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKGSEAGPAAISIIIAQAECECSQESPTFSERIFIAGSKQYSQESLQIPNN 120
Db 61 VITKGTAKGSEAGPAAISIIIAQAECECSQESPTFSERIFIAGSKQYSQESLQIPNN 120
QY 121 VAHATGKWARCVKGRSKARKRKXSSKSLAHAGVALAKPLPRTEQESCTIPVQE 180
Db 121 VAHATGKWARCVKGRSKARKRKXSSKSLAHAGVALAKPLPRTEQESCTIPVQE 180
QY 181 DESPLGAPVVRNTPQTKLPEGLGQLCFKQLEGRLPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPVVRNTPQTKLPEGLGQLCFKQLEGRLPRSELHKLISPLQCLNHVW 240
QY 241 KLHPDQGGPLPLTHPPFYSRLPHPPHPLQPKWPHPLESLGKLACVDSQKPLPDHP 300
Db 241 KLHPDQGGPLPLTHPPFYSRLPHPPHPLQPKWPHPLESLGKLACVDSQKPLPDHP 300
QY 301 LSKLACVDSQKPLPGHLPSCLSGAHEKFSVEEVLVHALOGSVSSQASHLTSIAKTW 360
Db 301 LSKLACVDSQKPLPGHLPSCLSGAHEKFSVEEVLVHALOGSVSSQASHLTSIAKTW 360
QY 361 AARGSRSEPSPKTEDNEGVLITKLPVDYREYREVEVHWATHQLRLGRGSFGEVHMDK 420
Db 361 AARGSRSEPSPKTEDNEGVLITKLPVDYREYREVEVHWATHQLRLGRGSFGEVHMDK 420
QY 421 QTGFQCAVKVRLVFRABELMACAGLTSPIRVLVYGVAREGPNWNI FWELEGSLGQL 480
Db 421 QTGFQCAVKVRLVFRABELMACAGLTSPIRVLVYGVAREGPNWNI FWELEGSLGQL 480
QY 481 VKEQGLCPEDRALYLGQALEGLEVLHSRIILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGLCPEDRALYLGQALEGLEVLHSRIILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGKSLLAGDYPGTETHMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600
Db 541 LQPDGLGKSLLAGDYPGTETHMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600
QY 601 RGPLCLKTASEPPVRETPPSCAPITQAQIQGLAKEPIHRSVAELGKGNRALQQVGG 660
Db 601 RGPLCLKTASEPPVRETPPSCAPITQAQIQGLAKEPIHRSVAELGKGNRALQQVGG 660
QY 661 LKSPWRGEYKEPRHPNPNQANVHOTLHAQPRELSRAPGPRPAETTTGRAPKQLPPLPE 720
Db 661 LKSPWRGEYKEPRHPNPNQANVHOTLHAQPRELSRAPGPRPAETTTGRAPKQLPPLPE 720
QY 721 PPEPNKSPPLTLKESGMEPLPLSSLEPAPARNPSPERKATVPEQLQLEIEFLN 780

Db 721 PPEPNKSPPLTLKESGMEPLPLSSLEPAPARNPSPERKATVPEQLQLEIEFLN 780
QY 781 SLSQPSFLEEQILSLCLSIDSLSDSEKKNPSKASOSSRDTLSSGVHWSQAEARS 840
Db 781 SLSQPSFLEEQILSLCLSIDSLSDSEKKNPSKASOSSRDTLSSGVHWSQAEARS 840
QY 841 SNMVLARGRPDTPSYFNGVKVQIQLNGEHLHIREFHRVKVGDIAIGTGISSQIPAAAFS 900
Db 841 SNMVLARGRPDTPSYFNGVKVQIQLNGEHLHIREFHRVKVGDIAIGTGISSQIPAAAFS 900
QY 901 LVTKDQGPVRYDMEVDSGIDLQCTLAPDGSFAWVRVKHGOLENRP 947
Db 901 LVTKDQGPVRYDMEVDSGIDLQCTLAPDGSFAWVRVKHGOLENRP 947
RESULT 2
Q8IYN1 PRELIMINARY; PRT; 947 AA.
AC Q8IYN1; OS Homo sapiens (Human);
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE MAP3K14 protein.
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RX MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RC SEQUENCE FROM N.A.
RP TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035576; AAH3576.1; --
DR HSSP; P71584; 106Y.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006488; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 947 AA; 104041 MW; C9D10F67FF7F49AC CRC64;

Query Match		99.8%; Score 5040; DB 2; Length 947;	
Best Local Similarity		99.8%; Pred. No. 2.5e-240;	
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1	MAVMEACPGAPGSAVGOKELPKPKETPLPKKQSSVYKLEAVEKSPVFCGWEILND	60
Db	1	MAVMEACPGAPGSAVGOKELPKAKETPLPKKQSSVYKLEAVEKSPVFCGWEILND	60
Qy	61	VITKGTAKGSEAGPAAISIIAQAECENSQFSTFSEIRIFIAQKQYQSSQESLDQIPNN	120
Db	61	VITKGTAKGSEAGPAAISIIAQAECENSQFSTFSEIRIFIAQKQYQSSQESLDQIPNN	120
Qy	121	VAHATEGMAVVCWKGKRSKARKKRSKSKSLAHAGVALAKPLPRTPEQESTTIPVQE	180
Db	121	VAHATEGMAVVCWKGKRSKARKKRSKSKSLAHAGVALAKPLPRTPEQESTTIPVQE	180
Qy	181	DESLGABYVNTQFTKPEGLGOLCFKQEGELRPAIPRSELKHLISPLQCLNHVW	240
Db	181	DESLGABYVNTQFTKPEGLGOLCFKQEGELRPAIPRSELKHLISPLQCLNHVW	240
Qy	241	KLHPQDGGPLPPTHPPYSLRPLPQPKPHLPESFLGKLACVDQSKPLDPDPH	300
Db	241	KLHPQDGGPLPPTHPPYSLRPLPQPKPHLPESFLGKLACVDQSKPLDPDPH	300
Qy	301	LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVAHQGVSSQASHLSLAKTW	360
Db	301	LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVAHQGVSSQASHLSLAKTW	360
Qy	361	AARSRSEPSKPTEDNEGVLLTEKLPVDYEEVEVHWATHQRLRGSGFGEVHRMEDK	420
Db	361	AARSRSEPSKPTEDNEGVLLTEKLPVDYEEVEVHWATHQRLRGSGFGEVHRMEDK	420
Qy	421	QTGFQCAVKVRLEVFRAELMACAGLTSPIVLYGAVREGPWNIIPMELLEGGSLQOL	480
Db	421	QTGFQCAVKVRLEVFRAELMACAGLTSPIVLYGAVREGPWNIIPMELLEGGSLQOL	480
Qy	481	VKEQGLPEDRALYLGQALEGLYLSRRILHGDVADNVLSSDGSAAALCDFGHAVC	540
Db	481	VKEQGLPEDRALYLGQALEGLYLSRRILHGDVADNVLSSDGSAAALCDFGHAVC	540
Qy	541	LQPDGLKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMHLMLNGCHPWTQFF	600
Db	541	LQPDGLKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMHLMLNGCHPWTQFF	600
Qy	601	RGPLCKIASPPVPREIPPCAPLTAQAIQEGLRKEPIHRVSAELGGKVNRLAQVGG	660
Db	601	RGPLCKIASPPVPREIPPCAPLTAQAIQEGLRKEPIHRVSAELGGKVNRLAQVGG	660
Qy	661	LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPRPABETTGAPKLPPLPPE	720
Db	661	LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPRPABETTGAPKLPPLPPE	720
Qy	721	PPBNKSPPLTLSKEESGWPEPLPLSSLEPAPARNPSPERKATVPPEOQLQELFLN	780
Db	721	PPBNKSPPLTLSKEESGWPEPLPLSSLEPAPARNPSPERKATVPPEOQLQELFLN	780
Qy	781	SLSPFSLSEOEQILCSLIDSLSDSDSEKNPKASQSSRDTLSSGVHSSQAEARSS	840
Db	781	SLSPFSLSEOEQILCSLIDSLSDSDSEKNPKASQSSRDTLSSGVHSSQAEARSS	840
Qy	841	SNNVLAGRGTPTSPYFNGVKVQIQSLNGSHLHREFHRVKVGDIGATIGSSQIPAAAFS	900
Db	841	SNNVLAGRGTPTSPYFNGVKVQIQSLNGSHLHREFHRVKVGDIGATIGSSQIPAAAFS	900
Qy	901	LVTXGQVRYDMEVPDSDIGDLQCTLAPDGSFAMSWRVKHQLENRP	947
Db	901	LVTXGQVRYDMEVPDSDIGDLQCTLAPDGSFAMSWRVKHQLENRP	947

Qy	16-OCT-2001 (Rel. 40, Created)	
DT	16-OCT-2001 (Rel. 40, Last sequence update)	
DT	05-JUL-2004 (Rel. 44, Last annotation update)	
DE	Mitogen-activated protein kinase kinase 14 (EC 2.7.1.37) (NF-kappa beta-inducing kinase) (Serine/threonine-protein kinase NIK).	
DE	Name=Map3k14; Synonyms=Nik;	
GN	Mus musculus (Mouse).	
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	SEQUENCE FROM N. A., AND VARIANT ALY ARG-855.
RP	STRAIN=C57BL/6; TISSUE=Spleen;	
RC	MEDLINE=99251583; PubMed=10319865; DOI=10.1038/9780;	
RX	Shinkura R., Kitada K., Matsuoka F., Tashiro K., Ikuta K., Suzuki M.,	
RA	Kogishi K., Serikawa T., Honjo T.;	
RA	"Alymphoplasia is caused by a point mutation in the mouse gene encoding Nf-kappa b-inducing kinase.";	
RT	Nat. Genet. 22:74-77(1999).	
RL	[2]	
RN	FUNCTION.	
RP	MEDLINE=21138222; PubMed=11239468; DOI=10.1016/S1097-2765(01)00187-3;	
RX	Xiao G., Hachaj E.W., Sun S.-C.;	
RA	"NF-kappaB-inducing kinase regulates the processing of NF-kappaB2	
RT	p100.";	
RL	Mol. Cell 7:401-409(2001).	
CC	-1- FUNCTION: Lymphotoxin beta-activated kinase which seems to be	
CC	exclusively involved in the activation of NF-kappa-B and its	
CC	transcriptional activity. Induces the processing of NF-kappa-B	
CC	2/P100. Could act in a receptor-selective manner.	
CC	-1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.	
CC	-1- SUBUNIT: Binds to TRAF2, TRAF3, TRAF6, IKKA and NF-kappa-B 2/P100.	
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	-1- PTM: Autophosphorylated.	
CC	-1- DISEASE: Defects in Map3k14 are the cause of alymphoplasia (ALY)	
CC	which is characterized by systemic absence of lymph nodes and	
CC	Peyer's patches and disorganized splenic and thymic structures	
CC	with immunodeficiency.	
CC	-1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP	
CC	kinase kinase subfamily.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; AF143094; AAD31512.1; -	
DR	MGD; MGI:1858204; Map3k14.	
DR	InterPro; IPR011009; Kinase like.	
DR	InterPro; IPR000719; Prot_kinase.	
DR	InterPro; IPR008271; Ser_thr_pkin_AS.	
DR	Pfam; PF00069; Pkinase; 1.	
DR	ProDom; PD000001; Prot_kinase; 1.	
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	
DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.	
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.	
KW	ATP-binding; Disease mutation; Phosphorylation;	
KW	Serine/threonine-protein kinase; Transferase.	
FT	DOMAIN 402 657 Protein kinase.	
FT	NP_BIND 408 416 ATP (By similarity).	
FT	BINDING 431 431 ATP (By similarity).	
FT	ACT_SITE 517 517 Proton acceptor (By similarity).	
FT	VARIANT 855 855 G -> R (in ALY; no binding to IKKA).	
SQ	SEQUENCE 942 AA; 103079 MW; 3BE4E4BA2D25C200 CRC64;	

RESULT 3

M3KE MOUSE

ID M3KE_MOUSE

STANDARD; PRT; 942 AA.

Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Query Match

Best Local Similarity

Matches 797; Conservative 53; Mismatches 90; Indels 9; Gaps 4;

83.1%; Score 4196.5; DB 1; Length 942;

84.0%; Pred. No. 9.3e-199;

[illegible]

DE	Hypothetical protein FLJ16568.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Esophageal;
RA	Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA	Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA	Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA	Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA	Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA	Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA	Zsuzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RL	Submitted (MAR-2004) to the ENBL/GenBank/DBJ databases.
DR	EMBL; AKL131438; BADI18584.1;
DR	GO; GO:0005524; F:ATP binding; IEA.
DR	GO; GO:0004672; P:protein kinase activity; IEA.
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR	InterPro; IPR011009; Kinase_like.
DR	InterPro; IPR000719; Prot_kinase.
DR	Pfam; PF00069; Pkinase; 1.
DR	ProDom; PD000001; Prot_kinase; 1.
DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW	Kinase.
SQ	SEQUENCE 477 AA; 52034 MW; ABBC068B4244B69E CRC64;
Query Match	
Best Local Similarity 45.5%; Score 2301; DB 2; Length 477;	
Matches 431; Conservative 1; Mismatches 2; Indels 0; Gaps 0	
QY	514 GDVADNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDYIRGTETHMAPEVVLRGSC 573
Db	44 GGLAADNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDYIRGTETHMAPEVVLRGSC 103
QY	574 DAKVDVWSSCCMLHLMNGCHPWTFQFRGPLCKIASEPPPVRIIPSCAPLTQAQIQEG 633
Db	104 DAKVDVWSSCCMLHLMNGCHPWTFQFRGPLCKIASEPPPVRIIPSCAPLTQAQIQEG 163
QY	634 LRKEPIHRVSAELGGKVNRALQVGIGLKFWRGYEKEPHRPPNQNYHTLHAQPREL 693
Db	164 LRKEPIHRVSAELGGKVNRALQVGIGLKFWRGYEKEPHRPPNQNYHTLHAQPREL 223
QY	694 SPRAGPRPAEETTRAPKLQPLPPPPPNNPKSPPLTLKSESGMWEPLPSLEPAPA 753
Db	224 SPRAGPRPAEETTRAPKLQPLPPPPPNNPKSPPLTLKSESGMWEPLPSLEPAPA 283
QY	754 RNSSSPRKATVPQEQLQQELEIELFNLSQLPSLEEQEILSCLSDLSDDSEKNP 813
Db	284 RNSSSPRKATVPQEQLQQELEIELFNLSQLPSLEEQEILSCLSDLSDDSEKNP 343
QY	814 SKASQSRRDTLSSGVHSWSCAARSSSNMVLARGPTDTPSYFNGVKVQIQSLNGEHL 873
Db	344 SKASQSRRDTLSSGVHSWSCAARSSSNMVLARGPTDTPSYFNGVKVQIQSLNGEHL 403
QY	874 HIREFHRVKVDGIATGISSQIPAAAFSLVTKGQPVRYDMVEVPDSGDIDLOCTLAPDGSA 933
Db	404 HIREFHRVKVDGIATGISSQIPAAAFSLVTKGQPVRYDMVEVPDSGDIDLOCTLAPDGSA 463
QY	934 WSWRVKHGLEENRP 947
Db	464 WSWRVKHGLEENRP 477
RESULT 5	
ID Q68D39	PRELIMINARY; PRT; 433 AA.
AC Q68D39	
DT 25-OCT-2004	(TrEMBLrel. 28, Created)
DT 25-OCT-2004	(TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004	(TrEMBLrel. 28, Last annotation update)
DE	Hypothetical protein DKFPz686J04131 (Fragment).

RESULT	4
Q6ZMW1	
ID	Q6ZMW1
AC	PRELIMINARY;
DT	PRT; 477 AA.
DT	05-JUL-2004
DT	(TRENMBrel. 27, Created)
DT	05-JUL-2004
DT	(TRENMBrel. 27, Last sequence update)
DT	05-JUL-2004
DT	(TRENMBrel. 27, Last annotation update)

RESULT 5	
Q68D39	
ID	Q68D39
AC	Q68D39;
DT	25-OCT-2004
DT	25-OCT-2004 (TRENBLrel. 28, Created)
DT	25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT	25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE	Hypothetical protein DKFZp686J04131 (Fragment).

```
GN Name=DKFZp686J04131;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RG The German cDNA Consortium;
RA Bloeker H., Boecker M., Brandt P., Mewes H.W., Weil B., Amid C.,
RL Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: CE749592; CAH18391.1;
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser Thr kinase.
DR InterPro: IPR008271; Ser Thr pkin_AS.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TK; 1.
DR SMART: SM00219; TyrK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 433 AA; 47220 MW; F9A8E975968AAOCA CRC64;

Query Match 41.8%; Score 2110; DB 2; Length 433;
Best Local Similarity 86.5%; Pred. No. 2.5e-96;
Matches 397; Conservative 3; Mismatches 15; Indels 44; Gaps 3;

QY 209 CFKQLGEGRLPALPRSELHKLISPLQCLNHWKLLHPDGGPLPHTPPFVSNLPHPFP 268
Db 1 CFKQLGEGRLPALPRSELHKLISPLQCLNHWKLLHPDGGPLPHTPPFVSNLPHPFP 268
QY 269 FHPQPWKPHLESFLGKLACVDSQKPLDPHLSKLACVDSQKPLGPHLEPSCLSRGAH 328
Db 52 FHPQPWKPHLESFLGKLACVDSQKPLDPHLSKLACVDSQKPLGPHLEPSCLSRGAH 328
QY 329 EKFSVEEYLVHVALQGSVSSQAHSLTSLAKTWAAGSRSPSPKTEDENGVLTLEKLP 388
Db 112 EKFSVEEYLVHVALQGSVSSQAHSLTSLAKTWAAGSRSPSPKTEDENGVLTLEKLP 388
QY 389 VDYEVEEVHATHQLRLGRGSGFGEVHRMEDKQTGFQCAVKVRLEVFRAEELMACAGLT 448
Db 172 VDYEVEEVHATHQLRLGRGSGFGEVHRMEDKQTGFQCAVKVRLEVFRAEELMACAGLT 448
QY 449 SPRIVPLYGAVREGPWVNI FMELEGGSLGQLVKEQGCLPEDRALYILGQALEGLELYLHS 508
Db 232 SPRIVPLYGAVREGPWVNI FMELEGGSLGQLVKEQGCLPEDRALYILGQALEGLELYLHS 508
QY 509 RRIHGDVKADNVLSSDGSAAALCDPGHACVCLQPDGLGKSLLTGDIYPTGTETMAPEV 568
Db 292 RRIHGDVKADNVLSSDGSAAALCDPGHACVCLQPDGLGKSLLTGDIYPTGTETMAPEV 568
QY 569 LGRSCDAKVDVWSSCCMMLHMLNGCHPWTFQFRGFLCLKIASPPVPIPPSCAPLTAQ 628
Db 352 LGRSCDAKVDVWSSCCMMLHMLNGCHPWTFQFRGFLCLKIASPPVPIPPSCAPLTAQ 628
QY 629 AIQEGRLKEPIHRSVAELGGKVRNALQQVGLKSPMRG 667
Db 394 GTQHQQGE-----LQPGQAAGWEG 415

RESULT 6
M3K8 HUMAN
ID M3K8 HUMAN STANDARD; PRT; 467 AA.
AC P41279; Q14275; Q9HCB1;
DT 01-FEB-1995 (Rel. 31, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
```

```
DE Mitogen-activated protein kinase kinase kinase 8 (EC 2.7.1.37) (COT
DE proto-oncogene serine/threonine-protein kinase) (C-COT) (Cancer Osaka
DE thyroid oncogene).
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone;
RG Chan A.M., Chedid M., McGovern E.S., Popescu N.C., Miki T.,
RA Aaronson S.A.;
RA "Expression cDNA cloning of a serine kinase transforming gene.";
RL Oncogene 8:1329-1333(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell;
RG Rieder M.J., Carrington D.P., da Ponte S.H., Hastings N.C.,
RA Ahearn M.O., Kuldane S.A., Rajkumar N., Toth E.J., Yi Q.,
RA Nickerson D.A.;
RA "SeattlesNPS. NHLBI HL66682 program for genomic applications, UW-
RT FHCR, Seattle, WA (URL: http://pga.gs.washington.edu).";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-112 FROM N.A.
RC TISSUE=T-cell;
RG MEDLINE=20469504; PubMed=10896655; DOI=10.1074/jbc.M000382200;
RA Sanchez-Gongora E., Liebana C., de Gregorio R., Ballester A.,
RA Calvo V., Perez-Jurado L., Alemany S.;
RA "COT kinase proto-oncogene expression in T cells: implication of the
RT JNK/SAPK signal transduction pathway in COT promoter activation.";
RL J. Biol. Chem. 275:31379-31386(2000).
RN [6]
RP CHARACTERIZATION.
RG MEDLINE=92019818; PubMed=1833717;
RA Aoki M., Akiyama T., Miyoshi J., Toyoshima K.;
RT "Identification and characterization of protein products of the cot
oncogene with serine kinase activity.";
RL Oncogene 6:1515-1519(1991).
CC -1- FUNCTION: Able to activate NF-kappa-B 1 by stimulating proteasome-
mediated proteolysis of NF-kappa-B 1/p105. Plays a role in the
cell cycle. The longer form of cot has some transforming activity,
although it is much weaker than the activated cot oncoprotein.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SUBUNIT: Interacts with NFKB1/p105.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- ALTERNATIVE PRODUCTS:
Event=Alternative initiation;
Comment=2 isoforms, 58 kDa (shown here) and 52 kDa, are produced
by alternative initiation;
CC -1- TISSUE SPECIFICITY: Expressed in several normal tissues and human
tumor-derived cell lines.
CC -1- DEVELOPMENTAL STAGE: The 58 kDa form is activated specifically
during the S and G2/M phases of the cell cycle.
CC -1- PTM: Autophosphorylated. The longer form undergoes phosphorylation
on Ser residues mainly, and the shorter form on both Ser and Thr
residues.
```


FT MUTAGEN 1381 1381 T->S: Reduced kinase activity and autophosphorylation.
 FT MUTAGEN 1381 1381 T->E: Loss of kinase activity and autophosphorylation.
 FT MUTAGEN 1381 1381 T->A: Falls to activate MAP2K1, MAP2K4, MAP2K7, CHUK and IKKB.
 FT MUTAGEN 1393 1393 T->S: Reduced kinase activity and autophosphorylation.
 FT MUTAGEN 1393 1393 T->E: Loss of kinase activity and autophosphorylation.
 FT MUTAGEN 1393 1393 T->A: Loss of kinase activity and autophosphorylation. Fails to activate MAP2K1, MAP2K4, MAP2K7, CHUK and IKKB.
 FT MUTAGEN 1394 1394 I->A: Loss of NF-kappa-B transcription factor activity and reduced ability to activate MAP2K1, MAP2K4, MAP2K7. No effect on AP-1 activity or activation of CHUK and IKKB. Loss of binding to IKKB.
 FT MUTAGEN 1396 1396 F->A: Loss of AP-1 and NF-kappa-B transcription factor activity. Reduced ability to activate MAP2K1, MAP2K4, MAP2K7, CHUK and IKKB.
 FT MUTAGEN 1397 1397 M->A: Loss of AP-1 and NF-kappa-B transcription factor activity. Reduced ability to activate MAP2K1, MAP2K4, MAP2K7, CHUK and IKKB.
 FT MUTAGEN 1401 1401 V->A: Loss of AP-1 and NF-kappa-B transcription factor activity. Reduced ability to activate MAP2K1, MAP2K4, MAP2K7, CHUK and IKKB.
 FT MUTAGEN 1402 1402 L->A: Loss of AP-1 transcription factor activity and reduced ability to activate CHUK and IKKB. No effect on NF-kappa-B activity or activation of MAP2K1, MAP2K4, MAP2K7. Loss of binding to MAP2K4.
 FT MUTAGEN 1403 1403 R->A: Loss of AP-1 transcription factor activity, no effect on NF-kappa-B activity.
 FT MUTAGEN 1404 1404 G->A: Loss of AP-1 and NF-kappa-B transcription factor activity.
 FT CONFLICT 30 39 GGALQGSGA -> ALQSGG (in Ref. 2).
 FT CONFLICT 103 103 Missing (in Ref. 2).
 FT CONFLICT 257 257 V -> E (in Ref. 2).
 FT CONFLICT 307 307 M -> V (in Ref. 2).
 FT CONFLICT 413 413 S -> C (in Ref. 2).
 FT CONFLICT 559 559 V -> A (in Ref. 2).
 FT CONFLICT 883 883 V -> L (in Ref. 3).
 FT CONFLICT 1467 1467 V -> L (in Ref. 3).
 FT SEQUENCE 1493 AA; 161288 MW; CA65C9B7703C6BF9 CRC64;
 Query Match 7.0%; Score 355.5; DB 1; Length 1493;
 Best Local Similarity 23.6%; Pred. No. 2.9e-09;
 Matches 157; Conservative 81; Mismatches 278; Indels 149; Gaps 24;
 QY 77 AISIITAQACENSQSFSTFSEIRIFAGSKQVSQESLDQIPNNVAHATEGMARVCWK 136
 Db 879 SLQAVAPTCSUENSLEHTVIREKTKGLSATRLSASEDISDLRAGVSGVLPSSTTTEQ 938
 QY 137 KR--RSKARKRKXKSKSLAHAGVALAKPLPRTPPEQESCTIPVQEDSPGAPYVNT 193
 Db 939 PKPAVQTKGRPHSQCLNSPLSHA--QLMFAPSPAPCSAPSPV--DIS-----KHR 986
 QY 194 PQFTPLKEPGLGQLCFKQLQEGRLAPLPRSELHKLISPLQ--CLNHVYKLLHHPQDGGPL 251
 Db 987 PQAFVPEKIPS-----ASPTQ-QKFSLQFQRCSEH-----RDSQL 1023
 QY 252 PLPTHFPYSLRHPFPPLQPKPHEPLGLKACVDSQKPLPDPLHSLKLA----- 305
 Db 1024 ---SPVFTQSRPPPSNTHRPKPSRPV-----GSTSKLGDATKSSMTLDLGS 1069
 QY 306 ---CVDS-----PKPLPGPH-----LEPSC-LSRGAHEKFSVEEVLVHALQGSVSS-- 347
 Db 1070 SRCDDSGGCGNGNAVIPSDTETVTPVEDKRLDVTNELNSSIEDLLEASMPSSDTTIVT 1129

QY 348 -----SQASHLTSIAKTWAARSRSR---EPSP 372
 Db 1130 FKSEVAVLSPEKAENDTYKDDVNNHKNQCKEKQWEAESEEAALAIAMAMSASQDALPITVPOL 1189
 QY 373 KTEDEGVLITKLP-----VDYEYREVVHATHQLRLGRGSGFVGRHMEKQTKGF 424
 Db 1190 QVENGEDIIIOQDTPELPGHTKAKQPYREDAEWLKGQ-QIGLGAFSSCYQADVGTT 1248
 QY 425 QCACVKV-----RLEVFA--EELMACAGLTPRIVPLYGAVREGPWNIIFMELLE 473
 Db 1249 LMAVQVTVVRNTSSEQEVEVEALREERIMGMHLNPNILRMGLATCEKSNYNLFIEWMA 1308
 QY 474 GSGSLQVKEQCLPEDRALYVLGOALSGLEYLHRSRRILHGDVADNVLLSSDGGHAALC 533
 Db 1309 GGSVAHLISKYGAFKESVVVINYTEQLRLGLSYLHENQIHRDVKGANLLIDSTGQRLRIA 1368
 QY 534 DFHGAVCLQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCMMHMLNGC 593
 Db 1369 DEGAARLASKGTGAGEFGQLL-GTIAFMAPEVLGRGOYGRSCDVWSVGCAIEMACAK 1427
 QY 594 HPWTOFFRPG---PLCLKTIASE-----PPPVREIPSPCAPLTAQ--AIQEGRLREP 638
 Db 1428 PPWNAEKHSNHLALIFKIASATTAPSIPLSHLSPGLRDVAVRCLELPQDRPSPRELLKHP 1487
 QY 639 IHRVS 643
 Db 1488 VFRIT 1492
 RESULT 12
 M3K2 MOUSE
 ID_M3K2 MOUSE STANDARD; PRT; 619 AA.
 AC Q61083;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Mitogen-activated protein kinase kinase 2 (EC 2.7.1.37)
 DE (MAPK/ERK kinase kinase 2) (MEK kinase 2) (MEK 2).
 GN Name=Map3K2; Synonyms=Mekk2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96214986; PubMed=8621389; DOI=10.1074/jbc.271.10.5361;
 RA Blank J.L., Gerwins P., Elliott E.M., Sather S., Johnson G.L.;
 RT "Molecular cloning of mitogen-activated protein/ERK kinase kinases
 (MEKK) 2 and 3. Regulation of sequential phosphorylation pathways
 involving mitogen-activated protein kinase and c-Jun kinase.";
 RL J. Biol. Chem. 271:5361-5368(1996).
 RN [2]
 RP FUNCTION, AUTOPHOSPHORYLATION, AND MUTAGENESIS OF PHB-571; ISO-573;
 RP THR-575; GLN-576; PRO-577; PRO-580; LEU-582; PRO-583 AND VAL-586.
 RX PubMed=12659851;
 RA Huang J., Tu Z., Lee F.S.;
 RT "Mutations in protein kinase subdomain X differentially affect MEKK2
 and MEKK1 activity.";
 RL Biochem. Biophys. Res. Commun. 303:532-540(2003).
 RN [3]
 RP INTERACTION WITH MAP2K5.
 RX PubMed=12912994;
 RA Nakamura K., Johnson G.L.;
 RT "Pb1 domains of MEKK2 and MEKK3 interact with the MEK5 Pb1 domain for
 activation of the ERK5 pathway.";
 RL J. Biol. Chem. 278:36989-36992(2003).
 CC -!- FUNCTION: Component of a protein kinase signal transduction
 cascade. Regulates the JNK and ERK5 pathways by phosphorylating
 and activating MAP2K5 and MAP2K7.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- COFACTOR: Magnesium.

CC -!- ENZYME REGULATION: Activated by phosphorylation on Thr-524 (By
CC similarity).
CC -!- SUBUNIT: Binds both upstream activators and downstream substrates
CC in multimeric complexes.
CC -!- PTM: Autophosphorylated.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
CC kinase kinase subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U43186; AAB03536.1; -;
CC HSSP; Q13153; 1F3M.
CC MGD; MGI:1346873; Map3k2.
CC InterPro; IPR011009; Kinase like.
CC InterPro; IPR000270; OPR_PBI.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_pkin_AS.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00564; PBI; 1.
CC Pfam; PF00069; Pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00666; PBI; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
CC ATP-binding; Magnesium; Metal-binding; Phosphorylation;
CC Serine/threonine-protein kinase; Transferase.
CC DOMAIN 356 616
CC NP_BIND 362 370
CC BINDING 385 385
CC ACT_SITE 483 483
CC MOD_RES 524 524
CC MUTAGEN 571 571
CC
CC MUTAGEN 573 573
CC
CC MUTAGEN 575 575
CC
CC MUTAGEN 576 576
CC
CC MUTAGEN 577 577
CC
CC MUTAGEN 580 580
CC MUTAGEN 582 582
CC
CC MUTAGEN 583 583
CC
CC MUTAGEN 586 586
CC SEQUENCE 619 AA; 69835 MW; 799A35218DDE7C41 CRC64;

Query Match 7.0%; Score 353; DB 1; Length 619;
Best Local Similarity 27.4%; Pred. No. 1.4e-09;
Matches 147; Conservative 64; Mismatches 201; Indels 124; Gaps 26;
QY 177 PQQSD--ESPLGAP-----YVRNTPQFTKPLKEG-----LGOLC-----EKQL-GEGLRP 219
DB 134 PSPDLNNTPLUGAERKRLSVGPPNDRSPPPGYIPDILHQIARNGSFTSINSE--E 191
QY 220 ALPRSELHLKLTSPQCLNHWKLHHPQDGG-----PLPLTPHPFYSLRPLPPFPFPHPL 272
DB 192 FIPES-WDQMLDPL-----SLSSPENSNGSCPSLDSPDGESYPKSRNPR----- 236
QY 273 QPMKHPHLESLFKLACVDSQKPLPDHLSKLACVDSKPLPGPHLEPSCLSRGAHEKFS 332
DB 237 -----AQSYPDNHQE-----FTDYDNPIFEKFGKGTYPRRYHVS 273
QY 333 VEEY-----LVHALQGS-----VSSQA-HSLTSLAKTWAARGSRSPSKPTD-- 376
DB 274 HQEYNDGRKTFPRARTQGTFSRSPSVFSPFDHSLST-----SSGSSVFTPEYDDSR 325
QY 377 -----NEGVLLTEKLPVDYREYREYVHWATHQLRLGRSGFGEVHRMDEKQTGFQCAV 428
DB 326 IRRGSDIDNPTLTVDISPPSRSPRAPTWRLGKL-LGQAGFGRVILCYDVTGRELAV 384
QY 429 KKVRL-----EVFRAE-ELMACAGLTSPRIPLYGAVR--EGPMVNIIFMELLEGS 476
DB 385 KQVQNPESPETSKEVNALECEIQLKNLLHERIVQYVCLURDPQKLTSLIFMELSPGS 444
QY 477 LGQLVKEQGCLPEDRALYIYLGQALEGLYLSRILHGDVKNADNVLLSSDSHAALCDFG 536
DB 445 IKDQLKAGALTENVTRKYTRQILEGVHYLSNMIVHRDIKGANILRDSTG-NIKLGDFG 503
QY 537 -----HAVCLQPDGLKSLTLTGDIPTGTHMAPEVVLGRSCDAKVDVWSSCCMMLHMLN 591
DB 504 ASKRLQICLSGTGM-KS-----VTGTPYMSFEVISEGEGYGRKADIWSVACTVEMLT 556
QY 592 GCHPWTQFFRGPLCLKIASEPPVREIPSPCAPTAQAIQEGLRKEPIHRVSAEL 647
DB 557 EKPWAEFEAMAAIFKIATQPTNPK-LPPHVSVDYTRDFLKR-IFVEAKLRPSAEL 610
RESULT 13
M3K1 RAT
ID M3K1 RAT STANDARD; PRT; 1493 AA.
AC Q62925; 1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Mitogen-activated protein kinase kinase 1 (EC 2.7.1.37)
DE (MAPK/ERK kinase 1) (MEK kinase 1) (MEK 1).
GN Name=Map3k1; Synonyms=Mekk, Mekki;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND MUTAGENESIS OF ASP-1369.
RC TISSUE=Brain;
RX MEDLINE=96224276; PubMed=8643568; DOI=10.1073/pnas.93.11.5291;
RA Xu S., Robbins D.J., Christerson L.B., English J.M., Vanderbilt C.A.,
RA Cobb M.H.;
RT "Cloning of rat MEK kinase 1 cDNA reveals an endogenous membrane-
RT associated 195-kDa protein with a large regulatory domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:5291-5295(1996).
CC -!- FUNCTION: Component of a protein kinase signal transduction
CC cascade. Activates the ERK and JNK kinase pathways by
CC phosphorylation of MAP2K1 and MAP2K4. Activates CHUK and IKKB,
CC the central protein kinases of the NF-kappa-B pathway (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- COFACTOR: Magnesium (By similarity).
CC -!- ENZYME REGULATION: Activated by phosphorylation on Thr-1381 and

Thr-1393 (By similarity).
 -|- SUBUNIT: Binds both upstream activators and downstream substrates in multimolecular complexes through its N-terminus (By similarity).
 -|- SUBCELLULAR LOCATION: Membrane-associated.
 -|- TISSUE SPECIFICITY: Most highly expressed in spleen, kidney and lung.
 -|- PTM: Autophosphorylated.
 -|- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP kinase kinase subfamily.
 -|- SIMILARITY: Contains 1 RING-type zinc finger.
 -|- SIMILARITY: Contains 1 SWIM-type zinc finger.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 EMBL; U48596; AAC52596.1; --
 PIR; T10757; T10757.
 KSSP; Q46539; IKV1.
 RGD; 620966; Map3K1.
 InterPro; IPR011009; Kinase like.
 InterPro; IPR000719; Prot kinase.
 InterPro; IPR008271; Ser Thr pkin_AS.
 InterPro; IPR002290; Ser-thr_pkinase.
 InterPro; IPR001841; Znf_ring.
 InterPro; IPR007527; Znf_SWIM.
 Pfam; PF00069; Pkinase; 1.
 Pfam; PF04434; SWIM; 1.
 ProDom; PD000001; Prot kinase; 1.
 SMART; SM00184; RING; 1.
 SMART; SM00220; S_TKC; 1.
 PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
 PROSITE; PS00011; PROTEIN KINASE_DOM; 1.
 PROSITE; PS00108; PROTEIN KINASE_ST; 1.
 PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
 PROSITE; PS00089; ZF_RING_2; 1.
 PROSITE; PS00966; ZF_SWIM; 1.
 ATP-binding; Magnesium; Metal-binding; Phosphorylation;
 Serine/threonine-protein kinase; Transferase; Zinc-finger.
 ZN_FING 328 356 SWIM-type.
 ZN_FING 433 482 RING-type.
 ZN_FING 1224 1489 Protein kinase.
 DOMAIN 25 29 Poly-Gly.
 DOMAIN 74 149 Pro-rich.
 DOMAIN 233 291 Pro-rich.
 DOMAIN 412 421 Poly-Ser.
 DOMAIN 1163 1168 Poly-Glu.
 NP_BIND 1230 1238 ATP (By similarity).
 BINDING 1253 1253 ATP (By similarity).
 ACT_SITE 1350 1350 Proton acceptor (By similarity).
 MOD_RES 1381 1381 Phosphothreonine (By similarity).
 MOD_RES 1393 1393 Phosphothreonine (By similarity).
 MUTAGEN 1369 1369 D->A: Inactivation.
 SEQUENCE 1493 AA; 161315 MW; 8C5P29F866898524 CRC64;
 Query Match 7.0%; Score 353; DB 1; Length 1493;
 Best Local Similarity 22.1%; Pred. No. 3.8e-09;
 Matches 165; Conservative 94; Mismatches 302; Indels 184; Gaps 25;
 QY 40 YKLEAVEKSPVPCGKWEILNDVITKGTAKGEAGPAALISITAAQACENSQFFS----- 93
 Db 791 FALQSIDNSHSMVGK---LSRRIYLSARWTVTPFLSKVTMLASGSSHFAKRRRL 847
 QY 94 PTFSEIRFIA-----GSK-----QYQSESLELDQIPNNVAHTEGKMARV 132
 Db 848 MAIADEVEIAEVIQLGSEDTLDGQDQDSQALAPPYPPESSSLE---HTAHVE----- 896
 QY 133 CWKGRKRKARKKRKKSSKSLAH--AGVALAKPLPRTPQEBSCTIPVQEDSESLGAPY- 189

Db 897 ----KTGKGLKATRLSASSEDISDLAGVSVGLPSSATTEQPKPTVQTK-----GRPHS 946
 QY 190 --VRNTPQTKPLKEPGLGQLCFKQGLRGLPALPRSELHKLISPLQCLNHVWKLHHPQD 247
 Db 947 QCLNSSPLSPQLMFFAISAPC-----SSAPSVAGSVTDA-----SKRPPA 989
 QY 248 GGPLPLPHTP-----FPYSRLPHPPFFHPLQPKPHP- 279
 Db 990 FVPCIKIPASPTQTKRFSLQFQTCSENRSSEKLSPVFTQSRPPSPSSNIHRAKASRPVPG 1049
 QY 280 LESFLGKLA-----CVDS-----QKPLPDHLSKLACVDSFKPLPGHL 318
 Db 1050 STSKLGDASKNSWTLDIANSQCDDSFSGNSGSAVIPSEETATPAEDKCRLDVNPDL 1109
 QY 319 -----EPCSLRGAHEKFSVEEYLVHALQSVSSQASHLS 355
 Db 1110 NSSIEDLLEASMPSSDITVTFKSEVAVLSPEKAESDDTYKDDVNHQCKEKMEAEDEEA 1169
 QY 356 LAKTWAARGSR---EPSPKTEDNEGVLLEKLP-----VDYREYREVHWATHQL 404
 Db 1170 LAIAWAMASQDALPIVPQLQVENGEDIIIIQQDTPETLPCHTKANEPYREDTEWLKQ- 1228
 QY 405 RLGRSGFGEVHMKDQTFQCAVKV-----RLEVFRA--EELMACAGLTSPRIV 453
 Db 1229 QIGLGFSSCYQAQDVGTGLMAVKQVTVYRNTSEQEEVVEALREERMHSLNHPNII 1288
 QY 454 PLYGNVREGPNWNIPELLEGGSLGQLVKEGCGLPEDRALYVLGQALEGLYLHRSRRLH 513
 Db 1289 RMLGATCEKSNYNLFIEWMAGSAVHLLSKYGAFKESVIVNVTQLLRGLSYLVHENQIIH 1348
 QY 514 GDVKADNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDIYIPCTETHMAPEVVLGRSC 573
 Db 1349 RDVKGANLLIDSTGQRLRIADFGAAARLASKGTGAGPEQQLL-GTIAFMAPEVLRGQY 1407
 QY 574 DAKVDVWSSCCWMLHMLNGCHPWTQFFRG---PCLCKIAS-----PPPVREIPP 620
 Db 1408 GRSCDVMSVGCALIEMACAKPPWNAEKSHNLALIFKIASATTAPSPHLSPLGRVAL 1467
 QY 621 SCAPLTAQ--AIQELGRKEPIHRVS 643
 Db 1468 RLELQPDQRPSPRELLKHPVFTT 1492
 RESULT 14
 Q6C2G1
 ID Q6C2G1 PRELIMINARY; PRT; 1158 AA.
 AC Q6C2G1;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Yarrowia lipolytica chromosome F of strain CLIB99 of Yarrowia
 DE lipolytica (fragment).
 GN ORFNames=YALI0F08165g;
 OS Yarrowia lipolytica CLIB99.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OX NCBI_taxID=284591;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB99;
 RG Genolevures;
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Laffontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
 RA Boismare A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
 RA Kerrest A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,
 RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
 RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,

Db 558 KPPWAEFEAMAAIFKTIATOPTNPK-LPPHVSDYTRDFLKR-IFVEAKLRPSADEL 610

Search completed: June 7, 2005, 12:02:21
Job time : 143.585 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:43:10 ; Search time 32.7215 Seconds
(without alignments)
2784.626 Million cell updates/sec

Title: US-09-155-676B-7
Perfect score: 5052
Sequence: 1 MAVMEMACPGAGSGAVGQQK.....PDGSFAWSRWKHCQLENRP 947

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	360	7.1	467	A48713	serine/threonine-s
2	357.5	7.1	467	I49609	proto-oncogene pro
3	356.5	7.1	467	A47388	serine/threonine p
4	355.5	7.0	687	A46212	MEK kinase - mouse
5	353	7.0	1493	T10757	MAP kinase kinase
6	336	6.7	1288	J80363	mitogen-activated
7	331	6.6	415	A41253	kinase-related tra
8	325	6.4	651	A96591	NPX1-related prote
9	325	6.4	1478	S20117	protein kinase BCK
10	319	6.3	653	T34356	hypothetical prote
11	313.5	6.2	883	A96662	hypothetical prote
12	307	6.1	696	S55694	protein kinase (EC
13	305.5	6.0	525	T58682	protein kinase, p2
14	305	6.0	550	T36746	probable serine/th
15	304	6.0	608	G96575	probable MEK kinas
16	303.5	6.0	1116	T38073	serine/threonine-p
17	301	6.0	706	T48084	STB1 protein kina
18	301	6.0	939	S28394	probable serine/th
19	298	5.9	1387	T16511	hypothetical prote
20	295	5.8	836	B96716	probable serine/th
21	294	5.8	1338	T30565	MAP kinase kinase
22	294	5.8	1401	T39225	MAP kinase kinase
23	292	5.8	544	S40482	serine/threonine-s
24	291.5	5.8	543	JC4070	protein kinase (EC
25	291.5	5.8	756	T50298	MAP kinase kinase
26	291	5.8	545	G01773	p21-activated prot
27	291	5.8	556	T36502	serine/threonine p
28	288.5	5.7	622	T15467	hypothetical prote
29	288	5.7	544	A57597	beta-p21-activated

30	287	5.7	1734	2	A54602	microtubule-associ
31	284	5.6	372	2	T52621	mitogen-activated
32	282	5.6	490	2	S47946	protein kinase hom
33	282	5.6	1062	2	S46367	protein kinase CDC
34	281	5.6	556	2	T42100	serine/threonine p
35	280	5.5	544	2	I49376	p21 activated kina
36	280	5.5	738	1	S51380	protein kinase STE
37	279	5.5	1230	2	T18256	probable serine/th
38	279	5.5	1230	2	T18259	serine/threonine p
39	279	5.5	1370	2	T19188	hypothetical prote
40	277.5	5.5	481	2	JE0377	p70 S6 kinase (EC
41	276.5	5.5	363	2	T51735	mitogen-activated
42	276.5	5.5	582	2	T51625	MAP3K alpha protei
43	276.5	5.5	626	2	B70754	probable serine/th
44	275	5.4	426	2	S71886	Ste20-like protein
45	275	5.4	1579	2	S59801	protein kinase SSK

ALIGNMENTS

RESULT 1

A48713

serine/threonine-specific protein kinase cot. 58K form - human

N:Alternate names: cot proto-oncogene, 58K form

N:Contains: serine/threonine-specific protein kinase cot, 52K form

C:Species: Homo sapiens (man)

C:Date: 07-Apr-1994 #sequence revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: A48713; B48713; S31639

R:Aoki, M.; Hamada, F.; Sugimoto, T.; Sumida, S.; Akiyama, T.; Toyoshima, K.

J. Biol. Chem. 268, 22723-22732, 1993

A:Title: The human cot proto-oncogene encodes two protein serine/threonine kinases with

A:Reference number: A48713; MUID:94043034; PMID:8226782

A:Accession: A48713

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-467 <AOK>

A:Cross-references: UNIPROT:P41279

A:Experimental source: TCO-4 cells

A>Note: sequence inconsistent with nucleotide translation

A>Note: sequence extracted from NCBI backbone (NCBIN:138970, NCBI:P138971)

A:Accession: B48713

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 30-467 <A02>

A:Experimental source: TCO-4 cells

A>Note: sequence inconsistent with nucleotide translation

A>Note: sequence extracted from NCBI backbone (NCBIN:138970, NCBI:P138972)

R:Chan, A.M.; Chedid, M.; Aaronson, S.A.; Miki, T.; McGovern, E.S.

submitted to the EMBL Data Library, July 1992

A:Description: A transforming gene isolated by expression cloning from Ewing's sarcoma c

A:Reference number: S31639

A:Accession: S31639

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-467 <CHA>

A:Cross-references: EMBL:Z14138; NID:G31244; PIDN:CAA78512.1; PID:G31245

C:Keywords: alternative initiators; proto-oncogene

F:136-388/Domain: protein kinase homology <KIN>

Query Match 7.1%; Score 360; DB 2; Length 467;

Best Local Similarity 27.2%; Pred. No. 1.5e-09;

Matches 116; Conservative 69; Mismatches 155; Indels 86; Gaps 14;

Qy 332 SVEYLVHVLQGSVSSQAHSLTSLAKTWAARGSRSPKTEDEGVLTKLKPVDY 391

Db 80 TVEDLL--AFANHISNTAKHY-----GORPOE-----SGILLNMTTPQNG 119

Qy 392 EYREE-----VHWATHQLRLQ-----RGSFGFVHRMEDKQTGFCVAKVKVLEVFRAEEL 441

Db 120 RYQIDSDVLLIPWKLTYRNTGSDFIPIRGAFKGYLAQDIKTKRMACKLIPVDQFKPSDV 179

Qy 442 MACAGLTSPIRVPLYGAVREGPWYNIPIMLLEGGSLGQLVKEQGCLPEDRALYYLGQALE 501

```
Db      180 EIQACFRHENIAELYGAVLWGTEVHLFWMEAGEGSVLEKLESCGPRREFELIIVWTKVLK 239
QY      502 GLEVLHSSRIILHGVDKADNVLLSSDGSAAALCDFGHAVCLQPDGLKSLTGYIP----- 557
Db      240 GLDLEHKKVHHIDIKPSNIVFMS--TKAVLVDFGLSVQMTED-----VYFFKDLR 288
QY      558 GTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTF--RGPLCLKIA-SEPP 613
Db      289 GTEIYMSPEVILCRGHTKADIYSLGATLIHMOTGTPPWKRYPRGAYPSYLYIHHKQAP 348
QY      614 PVREIPSCAPLTAQAIQEGRLKEPIHRVSAALGGKVNRAALQOVGLKSPMRGEYKEPR 673
Db      349 PLEDIADDCSPGMRELIEASLERNPNRPRAADL-----LK----- 384
QY      674 HPPPNQANYHOTLHAQPRELSPRAPGPRPAEETTGAPKLOPLPPEPPPEPKSPPLTLIS 733
Db      385 -----HEALN-PPREDQPRC---QSLDSALLERKLLSRKELELPENTADSSCTGS 431
QY      734 KERSGM 739
Db      432 TESEEM 437

RESULT 2
149609
proto-oncogene protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I49609
R;Ohara, R.; Miyoshi, J.; Aoki, M.; Toyoshima, K.
Jpn. J. Cancer Res. 84, 518-525, 1993
A;Title: The murine cot proto-oncogene 1: Genome structure and tissue-specific expression
A;Reference number: I49609; MUID:93308016; PMID:8320169
A;Accession: I49609
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-467 <RES>
A;Cross-references: UNIPROT:Q07174; GB:D13759; NID:G286072; PIDN:BAA02905.1; PID:G286073
C;Genetic:
A;Gene: cot
F;136-388/Domain: protein kinase homology <KIN>

Query Match          7.1%; Score 357.5; DB 2; Length 467;
Best Local Similarity 29.2%; Pred. No. 1.9e-09;
Matches 98; Conservative 60; Mismatches 115; Indels 63; Gaps 10;

QY      379 GVLLTEKLKPDVDEYREE-----VHWATHQLRLG-----RGSFGVHRMEDKQTGFQCAV 428
Db      107 GILLNNVISPQNGRYQIDSDVLLVPWKLTYRISIGSGFVPRGAFGKVIYLAQDMKTKRWAC 166
QY      429 KKVRLVFRABELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGGSLGQLVKEQGCLP 488
Db      167 KLIPIDQFKPSDVEIQACFRHENIAELYGAVLWGDTVHLFWMEAGEGSVLEKLESCGPMR 226
QY      489 EDRALYVLGQALEGLEYLHSRILHGDVKNVLLSSDGSAAALCDFGHAVCLQPDGLGK 548
Db      227 EFEIIVWTKHILKGLDFLHKKVHHIDIKPSNIVFMS--TKAVLVDFGLSVQMTED----- 280
QY      549 SLLTGDYIP---GTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTF--R 601
Db      281 -----VYLPKDLRGTEIYMSPEVILCRGHTKADIYSLGATLIHMOTGTPPWKRYPRSA 335
QY      602 GPLCLKIA-SEPPVPVREIPSCAPLTAQAIQEGRLKEPIHRVSAALGGKVNRAALQOVGG 660
Db      336 YPSYLYIHHKQAPPLEDIAGDCSPGMRELIEAALERNPNRPKAADL----- 382
QY      661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPR 696
Db      383 LK-----HEALN-PPREDQPR 397

RESULT 3
149609
proto-oncogene protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I49609
R;Ohara, R.; Miyoshi, J.; Aoki, M.; Toyoshima, K.
Jpn. J. Cancer Res. 84, 518-525, 1993
A;Title: The murine cot proto-oncogene 1: Genome structure and tissue-specific expression
A;Reference number: I49609; MUID:93308016; PMID:8320169
A;Accession: I49609
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-467 <RES>
A;Cross-references: UNIPROT:Q07174; GB:D13759; NID:G286072; PIDN:BAA02905.1; PID:G286073
C;Genetic:
A;Gene: cot
F;136-388/Domain: protein kinase homology <KIN>

Query Match          7.1%; Score 357.5; DB 2; Length 467;
Best Local Similarity 29.2%; Pred. No. 1.9e-09;
Matches 98; Conservative 60; Mismatches 115; Indels 63; Gaps 10;

QY      379 GVLLTEKLKPDVDEYREE-----VHWATHQLRLG-----RGSFGVHRMEDKQTGFQCAV 428
Db      107 GILLNNVISPQNGRYQIDSDVLLVPWKLTYRISIGSGFVPRGAFGKVIYLAQDMKTKRWAC 166
QY      429 KKVRLVFRABELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGGSLGQLVKEQGCLP 488
Db      167 KLIPIDQFKPSDVEIQACFRHENIAELYGAVLWGDTVHLFWMEAGEGSVLEKLESCGPMR 226
QY      489 EDRALYVLGQALEGLEYLHSRILHGDVKNVLLSSDGSAAALCDFGHAVCLQPDGLGK 548
Db      227 EFEIIVWTKHILKGLDFLHKKVHHIDIKPSNIVFMS--TKAVLVDFGLSVQMTED----- 280
QY      549 SLLTGDYIP---GTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTF--R 601
Db      281 -----VYLPKDLRGTEIYMSPEVILCRGHTKADIYSLGATLIHMOTGTPPWKRYPRSA 335
QY      602 GPLCLKIA-SEPPVPVREIPSCAPLTAQAIQEGRLKEPIHRVSAALGGKVNRAALQOVGG 660
Db      336 YPSYLYIHHKQAPPLEDIAGDCSPGMRELIEAALERNPNRPKAADL----- 382
QY      661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPR 696
Db      383 LK-----HEALN-PPREDQPR 397
```

```
A47388
serine/threonine protein kinase - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A47388
R;Patriotis, C.; Makris, A.; Bear, S.E.; Tschlis, P.N.
Proc. Natl. Acad. Sci. U.S.A. 90, 2251-2255, 1993
A;Title: Tumor progression locus 2 (Tpi-2) encodes a protein kinase involved in the pro-
A;Reference number: A47388; MUID:93211939; PMID:7681591
A;Accession: A47388
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-467 <PFI>
A;Cross-references: UNIPROT:Q63562; GB:M94454; NID:G207082; PIDN:AAA42185.1; PID:G207083
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBIN:128133, NCBI:P:128134)
F;136-388/Domain: protein kinase homology <KIN>

Query Match          7.1%; Score 356.5; DB 2; Length 467;
Best Local Similarity 29.2%; Pred. No. 2.1e-09;
Matches 98; Conservative 60; Mismatches 115; Indels 63; Gaps 10;

QY      379 GVLLTEKLKPDVDEYREE-----VHWATHQLRLG-----RGSFGVHRMEDKQTGFQCAV 428
Db      107 GILLNNVISPQNGRYQIDSDVLLVPWKLTYRISIGSGFVPRGAFGKVIYLAQDMKTKRWAC 166
QY      429 KKVRLVFRABELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGGSLGQLVKEQGCLP 488
Db      167 KLIPVDQFKPSDVEIQACFRHENIAELYGAVLWGDTVHLFWMEAGEGSVLEKLESCGPMR 226
QY      489 EDRALYVLGQALEGLEYLHSRILHGDVKNVLLSSDGSAAALCDFGHAVCLQPDGLGK 548
Db      227 EFEIIVWTKHILKGLDFLHKKVHHIDIKPSNIVFMS--TKAVLVDFGLSVQMTED----- 280
QY      549 SLLTGDYIP---GTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTF--R 601
Db      281 -----VYLPKDLRGTEIYMSPEVILCRGHTKADIYSLGATLIHMOTGTPPWKRYPRSA 335
QY      602 GPLCLKIA-SEPPVPVREIPSCAPLTAQAIQEGRLKEPIHRVSAALGGKVNRAALQOVGG 660
Db      336 YPSYLYIHHKQAPPLEDIAGDCSPGMRELIEAALERNPNRPKAADL----- 382
QY      661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPR 696
Db      383 LK-----HEALN-PPREDQPR 397

RESULT 4
A46212
MEK kinase - mouse
C:Species: Mus musculus (house mouse)
C>Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 28-May-1999
C:Accession: A46212
R;Lange-Carter, C.A.; Pleiman, C.M.; Gardner, A.M.; Blumer, K.J.; Johnson, G.L.
Science 260, 315-319, 1993
A;Title: A divergence in the MAP kinase regulatory network defined by MEK kinase and Rat
A;Reference number: A46212; MUID:93227040; PMID:8385802
A;Accession: A46212
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-687 <LAN>
A;Experimental source: brain
A;Note: sequence extracted from NCBI backbone (NCBI:P:129292)
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
C;Keywords: ATP
F;416-683/Domain: protein kinase homology <KIN>
F;424-432/Region: protein kinase ATP-binding motif

Query Match          7.0%; Score 355.5; DB 2; Length 687;
Best Local Similarity 23.6%; Pred. No. 3.5e-09;
Matches 157; Conservative 81; Mismatches 278; Indels 149; Gaps 24;

QY      77 AISIIAQECENSQSFSTPFSRIFTAGSKQYSQSESLDQIPNNVAHATEGKMARVCWK 136
```


Db 73 SLQVAPTSCLNSLEHTVIREKTGKLSATRLSASEDISDLRAGVSVGLPSTTTEQ 132
Qy 137 KR-----RSKARKRKKSLSLAHAGVALAKPLRTPPEQESCTIPVQEDSESLGAPYVRNT 193
Db 133 PKPAVOTKGRPHSQCNSPLSHA--QLMFPAPAPCSSAPSPV-----DIS-----KHR 180
Qy 194 POFKPLKEPLGQLCFKQGLGRLPALPRSELHKLISPLQ--CLNHWKLLHHPDQGGPL 251
Db 181 PQAEPVCKIPS-----ASPQTQ-RKFSLQFORNCSEH-----RDSQDL 217
Qy 252 PLPHPPYSRLPHFPPLPQPKPHEFLKGLACVDSQKPLPPHLSKLA----- 305
Db 218 ---SPVTQSRPPSSNIHRPKSRPV-----GTSKLGATKSMTLDLGS 263
Qy 306 ---CVDS-----PKPLPGPH-----LEPSC-LSRGAHEKFSVEEYLVAHALQSVSS-- 347
Db 264 SRCDDSGGGGNSGNAVIPSDETPTVEDKCRLDVNTLNSIEDLLEASMPSSDITVT 323
Qy 348 -----SQASHLTSIAKTWAARGSR-----BPSP 372
Db 324 FKSEVAVLSPEKAENDDTYKDDVNVHNOCKEKEAEAEALAIAMAMASQDALPIVPQL 383
Qy 373 KTEDNEGVLTEKLP-----VDYEYREEVHATHOLRLGRSGEVHRMEDKOTGP 424
Db 384 QVENGEIIIIQDTPETLPHTKAKOPYREDAEWLKGQ-QIGLGAFSSCYQAQDVGTGT 442
Qy 425 QCAVKV-----RLEVFA--EELMACAGLTSPIVLYGAVREGPWNIIFMELLE 473
Db 443 LMAVKQTYVNTSSEQEEVVEALREIRMMHNLNHNPIRLMGLATCEKSNYNLFIEMMA 502
Qy 474 GGSGLQVKEGCLPEDRALYYLQALLEGYLHSLRILHGDVKADNVLLSSDGSAAALC 533
Db 503 GGSVAHLLSKYGAFTSVINYTEQLLRGLSYLHENQIHRDVKGANLLIDSTQRLRIA 562
Qy 534 DFHAVCLQDGLGKSLTGDYIPGTETHMAPEVILGRSCDAKVDVWSSCCMLHMLNGC 593
Db 563 DFGAARLASKGTGAGBFQGLL-GTIAFMAPEVILGRSGQVGRSCDVWVSGCAIEMACAK 621
Qy 594 HPWTQFRG---PLCLKIASB-----PPVREIPSPCAPLTAQ--AIQGLRKEP 638
Db 622 PPWNAEKSHNLALIFKIASATTAPSPHLSPLGRDVAVRCLQLQDPPPSRELLKHP 681
Qy 639 IHRVS 643
Db 682 VFRTT 686

RESULT 5

Tl0757
MAP kinase kinase (EC 2.7.1.1) 1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Aug-2004
C:Accession: Tl0757
R;Xu, S.; Robbins, D.J.; Christerson, L.B.; English, J.M.; Vanderbilt, C.A.; Cobb, M.H.
Proc. Natl. Acad. Sci. U.S.A. 93, 5291-5295, 1996
A:Title: Cloning of rat MEK kinase 1 cDNA reveals an endogenous membrane-associated 195-
A:Reference number: 217123; MUID:96224276; PMID:8643568
A:Accession: Tl0757
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1493 <XUS>
A:Cross-references: UNIPROT:Q62925; EMBL:U48596; NID:gl354136; PIDN:AA52596.1; PID:gl35
C:Genetics:
A:Gene: MEK1
C:Superfamily: protein kinase homology
C:Keywords: ATP; phosphotransferase; protein kinase

Query Match 7.0%; Score 353; DB 2; Length 1493;
Best Local Similarity 22.1%; Pred. No. 1e-08;
Matches 165; Conservative 94; Mismatches 302; Indels 184; Gaps 25;

Qy 40 YKLEAVEKSPVFCGKWEILNDVITKTAKGESEAGPAAISIIAQACENSQEFSS----- 93

Db 791 FALQSDINSHMWGK---LSRRIYSSARMVTVPPPLFSKLVTWLSASGSHFARMRRRL 847
Qy 94 PTFSEIRIFIA-----GSK-----QYQSSESLDQIPNNVAHATEGQWARV 132
Db 848 MAIADAEVIAEVIQLGSEDLTGQDSSQALAPPYPESPSSLE-----HTAHVE----- 896
Qy 133 CWGKRKSKARKKKSLSLAH--AGVALAKPLRTPPEQESCTIPVQEDSESLGAPY- 189
Db 897 ----KTGKGLKATRLSASEDISDLRAGVSVGLPSSATTEQPKPTVQTK-----GRPHS 946
Qy 190 --VNTPTQFTKPLKEPLGQLCFKQGLGRLPALPRSELHKLISPLQCLNHWKLLHHPDQ 247
Db 947 QCLNSSLPSLPQMFPAISAPC-----SSAPSVAGSVTDA-----SKHRPRA 989
Qy 248 GGPLPLTPH-----FPYSRLPHFPFPHLPQWPKHP- 279
Db 990 FVPCIKFISAPQTKRKFSLQFQRTCSERNRDEKLSPVFTQSRPPSSNIHRAKASRPVPG 1049
Qy 280 LESPLGLKA-----CVDS-----QKPLPDPHLSKLKACVDSPKPLPGPHL 318
Db 1050 STSKLGDASKNSMTLDLNSASQCDDSPGSGNSGSAVIPSEETAFTPAEDKCRLDVNPDL 1109
Qy 319 -----EPSCLSRGAHEKFSVEEYLVAHALQSVSSQAHSLS 355
Db 1110 NSSIEDLLEASMPSSDPTVTFKSEVAVLSPEKAESDDTYKDDVNVHNOCKEKEAEAE 1169
Qy 356 LAKTWAARGSR-----BPSPKTEDNEGVLTEKLP-----VDYEYREEVHATHQL 404
Db 1170 LATAMASASQDALPIVPQLQVENGEDIIIIIQDTPETLPHTKANEPYREDTEWLKQ- 1228
Qy 405 RLGRSGEVRHMDKOTFCQAVKV-----RLEVFA--EELMACAGLTSPIV 453
Db 1229 QIGLGAFSSCYQAQDVGTGLMAVKQTYVNTSSEQEEVVEALREIRMMHNLNHNPII 1288
Qy 454 PLYGAVREGPWNIIFMELLEGGSLQGLVKEGCLPEDRALYYLQALLEGYLHSLRRLI 513
Db 1289 RMLGATCEKSNYNLFIEMMAGASVAHLLSKYGAFTSVINYTEQLLRGLSYLHENQI 1348
Qy 514 GDVKADNVLLSSDGSAAALCDFGHAVCLQDGLGKSLTGDYIPGTETHMAPEVILGRSC 573
Db 1349 RDVKGANLLIDSTQRLRIAIDFGAARLASKGTGAGBFQGLL-GTIAFMAPEVILGRQY 1407
Qy 574 DAKVDWSSCCMLHMLNGCHPWTQFRG---PLCLKIASB-----PPVREIPSP 620
Db 1408 GRSCDVWVSGCAIEMACAKPPWNAEKSHNLALIFKIASATTAPSPHLSPLGRDVAL 1467
Qy 621 SCAPLTAQ--AIQGLRKEPIHRVS 643
Db 1468 RCLELOQDPPPSRELLKHPVFRIT 1492

RESULT 6

JE0363
mitogen-activated protein kinase kinase kinase (EC 2.7.-.-) - human
C:Species: Homo sapiens (man)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C:Accession: JE0363
R;Wang, X.; Diener, K.; Tan, T.; Yao, Z.
Biochem. Biophys. Res. Commun. 253, 33-37, 1998
A:Title: MAPKKK6, a novel mitogen-activated protein kinase kinase kinase, that associate
A:Reference number: JE0363; MUID:99092374; PMID:9875215
A:Accession: JE0363
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1288 <WAN>
A:Cross-references: GB:AF100318
C:Keywords: phosphotransferase

Query Match 6.7%; Score 336; DB 2; Length 1288;
Best Local Similarity 24.0%; Pred. No. 5.2e-08;
Matches 177; Conservative 92; Mismatches 276; Indels 194; Gaps 31;

```
QY 149 KSSKSLAHAGVALAKPLPRTPQESCTIPQV---EDESPGLGAVYVRNTPTQTKPLKEPGL 205
Db 410 EDSKELRLIGKGLCLLAR-----KGCVERKQYWDVGVYGAQLLNDP-----454
QY 206 GQLCFKQGLGELRPALELHKLISPLQCLNHVK---LHHQDGGPLPLTHPPPYSR 262
Db 455 -----TQVVLAAEQLYKLNAPIWLVSVMTFLY--QHGRPTPEP--PGGPPR 499
QY 263 LPHPPFPHPQPKPPLPGLKACVDSQKPLPDPHLSKLACVSPKPLPGPHLEPSC 322
Db 500 RAHFWLHFLQSQQPF-----KTAQAQGDQCL-----VLVLENNKVLPAKLE---542
QY 323 LSRGAHEKFSVEEYLVHALQSVSSQAHSLSLTAQWAAAGSR-----366
Db 543 -VRGTDPSVSTVLSLEPEQDIPSSWTFPVAICGVASAKRDERCCFLYALPPAODVOL 601
QY 367 -----SRPSPKTEDEGVLLETKLPVDVYEVREVEVHATHQLRL 406
Db 602 CFFSVGHCFWCGLIQAWVTNPDSTAPAEAEAG---AEEMLEPDYEYTE---TGERLVL 654
QY 407 GRGSFGFVHRMEDKQTFQCAVKV-----RLEVFRAEELMACAGLTSRIVPLYCAVRE 461
Db 655 GKTYGVVYAGRDHRTVRIKIEIPERDSRFQPLHEEIALHRRHKNIVRYLGSASQ 714
QY 462 GPWNIPMELLEGSLGQVKE--QGCLPEDRAL--YYLQGALEGLYLEHSRRLHGDVKA 518
Db 715 GGYLKIFMEIEVPGCSLSLLRSVMGPKDNESITISFYTRIIQGLGYLDHNIHVHDIKG 774
QY 519 DNVLLSSDGSAAALCDFGHAVCLQPDGLKSLITG-----DYIPGTETHMAPEVV--LGR 571
Db 775 DNVLIINTFGSLKISDFGTS-----KRLAGITPCTETPTGTLOYMAPEIIDQGR 824
QY 572 SCDKVDVWSSCCMMLHMLNGCHPWTOFFRGPLCLKIASEPPPVREI-----PPSCAPL 625
Db 825 GYCKADINSLGCTVIEMATGRPFHE-----LGSPPQAMFGVMGVKHPMPWSSL 875
QY 626 TAQA-----IOEGLRKEPIHRVSAEELGGKVNALQQVGGGLKSPWRGEYKEPRHPPPNQANY 682
Db 876 SAEAQAFLRTPEPDLRASATLLG--DPFLQPGKRSRP-----SSPRHAP-----922
QY 683 HQTLLAQPLESPRAGPRPAETT-----GRAPKLOPPLPPE-----PP 722
Db 923 -----RPDASASSTP--SANSSTQSFPCQAPSOHPSPPPKCKSYGTSQQLRVPE 975
QY 723 EPNKSPPLTSLKEESGMWELPLSLLEPAPARNPSPERKATVPQELQQLIEFLNLSL 782
Db 976 EPAAEERPAS--PESSG-----LSLLHQESKRR-----AMLAVALQELPALAENLHQEQK 1024
QY 783 SQP---FSLEEOEQILSCL 798
Db 1025 QEQGARLGRNHVLELLRCL 1043

RESULT 7
A41253
kinase-related transforming protein (EC 2.7.1.-) - human
C:Species: Homo sapiens (man)
C>Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 15-Mar-2004
C:Accession: A41253
R:Miyooshi, J.; Higashi, T.; Mukai, H.; Ohuchi, T.; Kakunaga, T.
Mol. Cell. Biol. 11, 4088-4096, 1991
A:Title: Structure and transforming potential of the human cot oncogene encoding a putat
A:Reference number: A41253; MUID:191304400; PMID:2072910
A:Accession: A41253
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-415 <MIV>
C:Keywords: phosphotransferase
F:136-388/Domain: protein kinase homology <KIN>

Query Match 6.6%; Score 331; DB 2; Length 415;
Best Local Similarity 27.7%; Pred. No. 2.7e-08;
Matches 104; Conservative 65; Mismatches 136; Indels 70; Gaps 13;
```

```
QY 332 SVEEYLVHALQSVSSQAHSLSLTAQWAAAGSRSPKTEDEGVLLETKLPVDY 391
Db 80 TVEDLL--AFANHISNTAKHFY-----GQRQPE-----SGILLANWITPQNG 119
QY 392 EYREE-----VHWATH---QLRLGRSGFVGHVHRMEDKQTFQCAVKKRVLEVFRAEEL 441
Db 120 RYQIDSDVLLIISNTAKHFYQRPQESGILLKVVYLAQDIKTKRWACKLIPVDQFKPSDV 179
QY 442 MACAGLTSRIVPLYGAVREGPWNNIFMELLEGSLGOLVKEQGCLPEDRALYYLQGALE 501
Db 180 ETQACFRHENTAEELGAVLWGETVHLFMEAGEGGSVLEKLSGCGPMRFEETIIVWTKHLK 239
QY 502 GLEYLHSRRLHGDVADNVLLSSDPGSHAAALCDFGHAVCLQPDGLKSLTGDYIP-----557
Db 240 GLDPLHKKVIHHDIKPSNIVFMS--TKAVLVDFGLSVQMTED-----VYFFKDLR 288
QY 558 GTETHMAPEVVILGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF---RGPLCLKIA--SEPP 613
Db 289 GTEIYMSPEVILCRGHSTKADIYSLGATLIHMQTGTPPVKRYPRSAVPSYLYIIHKQAP 348
QY 614 PVREIPSPCAPLTAQAIQIEGLRKEPIHRVSAEELGGKVNALQQVGGGLKSPWRGEYKEPR 673
Db 349 PLEDIADDCSPGMRELIEASLERPNHRPRAUL-----LK-----HEAL 388
QY 674 HPP-PNQANYHQTLH 687
Db 389 NPPREDQPRGHQVIH 403

RESULT 8
A96591
NPKI-related protein kinase 2 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A96591
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,
Nure, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A96591
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-651 <STO>
C:Cross-references: UNIPROT:Q9FZ36; GB:AEO05173; NID:99857521; PIDN:AAG00876.1; GSPDB:GN
C:Genetics:
A:Gene: T24C10.7
A:Map position: 1
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

Query Match 6.4%; Score 325; DB 2; Length 651;
Best Local Similarity 26.1%; Pred. No. 8e-08;
Matches 132; Conservative 69; Mismatches 188; Indels 116; Gaps 18;

QY 383 TEKLKPYDVEYRBEVHWATHQLRLGRSGFGEVHRMEDKQTFQCAVKKRVL-----EV 435
Db 59 TVQIKP-----PIRWKQQL-IGRGAFGVYVMGNLSDGELLANKVQLITSNCASKEK 110
QY 436 FRA-----BELMACAGLTSRIVPLYGAVREGPWNNIFMELLEGSLGOLVKEQGCLPE 489
Db 111 TQAHIOELEBEVKLLKNLSHPNIVRYLGTVREDETINILIEFVPGGSISLLEKFGAPPE 170
QY 490 DRALYYLGOALEGLYLEHSRRLHGDVADNVLLSSDGSAAALCDFG---HAVCLQPDGL 546
Db 171 SVVRYTNQLLLGLEVYLNHAIHWRDIKGNILVDNQGCK-IKLADFGASKQVLAETISG 229
```

Qy 547 GKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMLHMLNGCHPWTOFFRGPLCL 606
Db 230 AKSM-----KGTPTWMAPEVILQTGHFSADINVSQCTVIEMVTGKAPWSQQYK----- 278
Qy 607 KIA-----SEPPVRETPSPCAPLTAQAIOEGLRAKEPHRVSAAELGKYNRALQQ 657
Db 279 EIAAFHGTYSKHP-----IPDNISSDANDFLLKLOEENLRTASEL----- 324
Qy 658 VGGLKSPW-RGEYKEPRHPNPNQANYHOTLHAQPRELSFRAPGRPAEETTRAPKLOPP 716
Db 325 ---LKHPPVTKQKE-----SASKDLTSFMDNSCSP 352
Qy 717 LPPEPPENKAPPLTSLKEBSQWEPLPLSLEPAPARNPSPERKATVPQEQLOOLEIE 776
Db 353 LPSE--LNTITSYQTSDDVG--DICNLGLTCTLA-----FPEKSIQNNSLC 397
Qy 777 LFLNSLQPFSELEEQILSLCLSIDLSLSDSEKNPS-----KASQSRDRTLSSGVHWS 832
Db 398 LKSNV-----GYDDDDNDMCLIDDENFLTYNGETGPSLDNNTDAKSCDTMSEISDLK 452
Qy 833 SQAEARSSNMVNLARGRPPTTPSY 857
Db 453 CKFDENSGNETETKVSMEVDHPSY 477

RESULT 9
S20117
protein kinase BCK1 (BC 2.7.1.1) - Yeast (Saccharomyces cerevisiae)
N:Alternate names: protein J0906; protein kinase SLK1; protein kinase SSP31; protein YU1
C:Species: Saccharomyces cerevisiae
C>Date: 23-Apr-1993 #sequence revision 23-Apr-1993 #text_change 16-Aug-2004
C:Accession: S20117; S22285; S19061; JQ1432; S56872; S30794; JQ1118
R:Costigan, C.; Gehring, S.; Snyder, M.
Mol. Cell. Biol. 12, 1162-1178, 1992
A:Title: A synthetic lethal screen identifies SLK1, a novel protein kinase homolog implied in cell proliferation
A:Reference number: S20117; MUID:92186847; PMID:1545797
A:Accession: S20117
A:Molecule type: DNA
A:Residues: 1-1478 <COS>
A:Cross-references: UNIPROT:Q01389; EMBL:M84389
A:Experimental source: strain S288C
R:Miosga, T.; Boles, E.; Schaff-Gerstenschlaeger, I.; Schmitt, S.; Zimmermann, F.K.
Yeast 10, 1481-1488, 1994
A:Title: Sequence and function analysis of a 9.74 kb fragment of Saccharomyces cerevisiae chromosome V
A:Reference number: S50295; MUID:95176706; PMID:7871887
A:Accession: S50298
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1478 <MIO>
A:Cross-references: EMBL:X77923; NID:9640004; PIDN:CAA54896.1; PID:9640009
R:Lee, K.S.; Levin, D.E.
Mol. Cell. Biol. 12, 172-182, 1992
A:Title: Dominant mutations in a gene encoding a putative protein kinase (BCK1) bypass the requirement for a second mutation in a gene encoding a putative protein kinase
A:Reference number: S22285; MUID:92107166; PMID:1729597
A:Accession: S22285
A:Molecule type: DNA
A:Residues: 1-58, 'I', 60-1478 <LBE>
A:Cross-references: EMBL:X60227
A:Experimental source: strain EG123
R:Lee, K.S.; Levin, D.E.
submitted to the EMBL Data Library, June 1991
A:Description: An extragenic suppressor of mutations in the S. cerevisiae protein kinase BCK1
A:Reference number: S19061
A:Accession: S19061
A:Molecule type: DNA
A:Residues: 1-58, 'I', 60-263, 'P', 265-278, 'I', 280-702, 'S', 707-708, 'KP', 714, 'VITWTE', 715-79
A:Cross-references: EMBL:X60227; NID:G3414; PIDN:CAA42788.1; PID:G3415
A:Experimental source: strain EG123
R:Irie, K.; Araki, H.; Oshima, Y.
Gene 108, 139-144, 1991
A:Title: A new protein kinase, SSP31, modulating the SMP3 gene-product involved in plasma membrane transport
A:Reference number: JQ1432; MUID:92104496; PMID:1840547
A:Accession: JQ1432

A:Molecule type: DNA
A:Residues: 149-1478 <IRI>
A:Cross-references: EMBL:D10389; DDBJ:D90446
R:Miosga, T.; Schaff-Gerstenschlaeger, I.; Baur, A.; Boles, E.; Chalwatzis, N.; Fournier, J.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56855
A:Accession: S56872
A:Molecule type: DNA
A:Residues: 1-1478 <MIW>
A:Cross-references: EMBL:Z49370; NID:G1008269; PIDN:CAA89389.1; PID:G1008270; MIPS:YJL095w
R:Cusick, M.E.
submitted to the EMBL Data Library, March 1992
A:Reference number: S27437
A:Accession: S30794
A:Molecule type: DNA
A:Residues: 602-959, 'R', 961, 'R', 963-1085, 'V', 1087, 'SLLIHT', 1092-1094, 'RMD', 1101, 'TV', 1111
A:Cross-references: EMBL:M86604; NID:G172073; PIDN:AAA21179.1; PID:G172074
C:Genetics: S27437
A:Gene: SGD:BCK1; SLK1; SSP31
A:Cross-references: SGD:S0003631; MIPS:YJL095w
A:Map position: 10L
C:Function:
A:Description: phosphotransferase; protein kinase; involved in cell proliferation
C:Superfamily: protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F:1173-1440/Domain: protein kinase homology <KIN>
F:1181-1189/Region: protein kinase ATP-binding motif

Query Match 6.4%; Score 325; DB 2; Length 1478;
Best Local Similarity 28.7%; Pred. No. 1.9e-07;
Matches 117; Conservative 53; Mismatches 163; Indels 74; Gaps 14;
Qy 258 FPYRLPHPPFPHLPQWPKHPLESFLGKACVDSQKPLPDPLHSLKACVDSQKPLPGPH 317
Db 1043 FPRANLQKPIGTEGASTSPKSLDLSL-----SPKNVASSR 1078
Qy 318 LEPSCLSRGAHEKFSVEYLVAHQGS---VSSQASHLSLAKTMAARGSRSPSPK- 373
Db 1079 TEPSTSPRPDPSSY-EFIQDGLGNKKNKLNQKPKTKTKTIRTIHAESLARKNSVKL 1137
Qy 374 -----TEDNEGVLITKLPVDYREVEVHWATHQLRLGRSGFGEVHRME 418
Db 1138 KRQTKMGTMTMVEVTNH---MVSINKAKNSKGYK-EFAMKGM-IGKSGFQAVYLC 1193
Qy 419 DKQTFQCAVKKVLVEFRAB-----ELMACAGLTPRIPLYGAVREGPW 465
Db 1194 NVTGEMVAKQVEVPKYSSQNEALSTVEALREVSTLKDLHLNIVQYLGFEKNKIY 1253
Qy 466 NIFMELGGSLGQVKEQGLPEDRALYVLGQALEGLYHLHSRRLHGDVKADNVLS 525
Db 1254 SLFLEYVAGSGVSLIRMYGRFDEPLIKHLTTQVLKGLAYLHSGILHRDMKADNLLDQ 1313
Qy 526 DGSAAALCDRGHACVCLQPDGLGKSLLTGDIYPTGTHWAPVW-LGRSCDAKVDVWSSCC 584
Db 1314 DGI-CKISDFGIS-RKSKDIYSNDSMT---MRGTVFWMAPEMVDTKQGSYSAKVDIWSLGC 1368
Qy 585 MMLHMLNGCHPWTOFFRGPLCLKIA---SEPPVRETPSPCAPLTAQ 628
Db 1369 IVLENFAGKRPWSNLEVVAMFKIGSKSAPP-----IPEDTLPLISQ 1411

RESULT 10
T34356
hypothetical protein T19A5.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T34356
R:Bradehaw, H.
submitted to the EMBL Data Library, March 1996
A:Description: The sequence of C. elegans cosmid T19A5.
A:Reference number: Z21512
A:Accession: T34356
A>Status: preliminary; translated from GB/EMBL/DBJ

Db 814 QDRSSRPTPHVFWNDNGSIQPCYNWN-----KDNQPVLSHDVSSQQLLS-----EHLK 863

QY 875 IREFHRVKVGIATGISS 892

Db 864 LKSL-----DLRPGFST 875

RESULT 12

S55694

protein kinase (EC 2.7.1.37) sk1, CAMP-dependent - fission yeast (Schizosaccharomyces p

C;Species: Schizosaccharomyces pombe

C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 16-Aug-2004

C;Accession: S55694; T38040

R;Jin, M.; Fujita, M.; Culley, B.M.; Apolinario, E.; Yamamoto, M.; Maundrell, K.; Hoffma

Genetics 140, 457-467, 1995

A;Title: sk1, a high copy number suppressor of defects in the CAMP-dependent protein ki

A;Reference number: S55694; MUID:96120227; PMID:7498728

A;Accession: S55694

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-696 <JIN>

A;Cross-references: UNIPROT:P50530; GB:D38108; NID:g1136301; PIDN:BAA07286.1; PID:d10078

A;Note: the authors translated the codon GAT for residue 687 as His

R;McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Whitehead, S.; Churcher, C.M.

submitted to the EMBL Data Library, August 1999

A;Reference number: Z21764

A;Accession: T38040

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-176, 'F', 178-198, 'A', 200-696 <MCD>

A;Cross-references: EMBL:AL109951; PIDN:CAB53053.1; GSPDB:GN00066; SPDB:SPAC1B9.02c

A;Experimental source: strain 972h; cosmid c1B9

C;Genetics:

A;Gene: sk1; SPDB:SPAC1B9.02c

A;Map position: 1

A;Introns: 80/3; 311/1; 633/2

C;Superfamily: protein kinase homology

C;Keywords: ATP; phosphotransferase; protein kinase

F;300-563/Domain: protein kinase homology <KIN>

F;308-316/Region: protein kinase ATP-binding motif

Query Match 6.1%; Score 307; DB 2; Length 696;

Best Local Similarity 21.5%; Pred. No. 5.6e-07;

Matches 187; Conservative 125; Mismatches 292; Indels 264; Gaps 42;

QY 36 QSSVYKLEAVEKSPVFCGKWEILNDVITKTAEGSEAGPAAISIIAQACENSQEFSP 95

Db 19 QASPTIQSHSTQPV-----LSNDHSTKVNDYEGEGA-----SSNGYDP- 58

QY 96 FSEIFTAGSKQYSQSESLDQIPNNVAHATEGKMARVCWKGKRSKARKKKKSSKSLA 155

Db 59 ----VFMSDRMKRYNEITAQ-----HKEQSLKDEKSGSNSSE 94

QY 156 HAGVALAKPLRTPPEQSCCTIPVQEDSPLGAPYVRNTPQTPLKPEGLQLCFKQL-- 213

Db 95 SNGIT--PMGTSEKPK-----LQSRTPPSSCVRHDTVPVKDKN---QGHAFGRLYV 143

QY 214 ----GEGRLPALPSELHKLIS-----PLQCLNHWKLLHHHPQDGG-PLPLPT--H 256

Db 144 RLHQGRDLNVKSVNAQPYAVITFEKTQVMVPPPLKDI-----DGGIPISPKNR 193

QY 257 PFPYSRLPFPFPFPLQPKPHLESFLGKLACVDSQKPLDPHLSKLACVDSPKPLPGP 316

Db 194 PLAGSRSSSGLHS-----ELMLADVRC-----PHWDETVPDVTK----- 230

QY 317 HLEPSCLSRGAHFKSPVEYLVHVALQGSVSSQASHLSLAKTWAAGSRGREPSKPTD 376

Db 231 --MKSQMVSVYDYEDDKFL-----GSVKITPFLFHEYVQEA-----YKLEPLDLTKS 278

QY 377 NEGVLLEKLPVYEVREYVHVAHQ-----LRL-GRGSFGVHVRNEKQTFQCAVKVR 432

Db 279 LEGBIKVETI-----YEHIEHVRYPEDFTALRLIGKTFQGVYLVKNDTNRIYAMKKIS 334

QY 433 LE-VFRAEELMACAGL-----TSPRIPLYGAVREGFWNIFFMELLEGSGLGOLV 481

Db 335 KKLIVRKKEVTHTLGERNILVRTSLDESPFTVGLKFSQFASDLYLITDYMGGSLFWHL 394

QY 482 KEQCLPEDRALYYLGOALEGLVLSRRILHGVKADNVLLSSDGSAAALCDFGHAVCL 541

Db 395 QHEGRFPEQRAKFVIAELVLALEHLKHDIIRDLKPNILLDDAG-HIALCDFG---L 449

QY 542 QPDGLGKSLTGDYIPGTETHAPEVVL-GRSCDAKVDMVSSCCMMLHMLGCHPEWTOFF 600

Db 450 SKANLSANATNTTC-GTTEVLAPEVLEDDKGYTKQVDFWSLGLVLFEM---CCGWSFFY 505

QY 601 RGPLCLKIASEPPVPREI-----PSCAPLPTAQAIQEG-LRKEPIHVRVSAEIG 648

Db 506 -----APDVQOQMYRIAFKRVFRKGVLSSEGRSFVRGLNPNHR-----LG 549

QY 649 GKVNRAQQVGLKSPWRGEYKEPRHPPPNPNQANYHTLHAQPRELSPRAGPRPAEETG 708

Db 550 AVADTT-----ELKE--HPFFADINWDL-----LSKK----- 574

QY 709 RAPKLQPLPPEPPPEPKSPPLTSL---KE--ESGMWEPLPLSLLEPAPARNPSSPERKA 763

Db 575 ---KVQPPF-----KPNVQNDLDVSNFDKFTNTNVKNINIVSNVDPANASTPLS----- 621

QY 764 TVPEQELQQLIELEFLNLSQPF-----SLEEQEQLSLCLSDLSLSDSSEKPNPK 815

Db 622 -----NTIQDRFRGFTFVNKSIDEQFQ-----NLGLQENETDNLH 657

QY 816 ASQSRDTLSSGVHSSWSQAEARSSWN 843

Db 658 A---CRTTHSSVNSINSNGNPRTVDAN 682

RESULT 13

S58682

protein kinase, p21-activated (EC 2.7.1.1) - human

N;Alternate names: protein kinase PAK65; S6/H4 kinase

C;Species: Homo sapiens (man)

C;Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 16-Aug-2004

C;Accession: S58682; S55258; S55304; S58690; A57441

R;Sells, M.; Knause, U.J.; Bagrodia, S.; Ambrose, D.; Bokoch, G.M.; Chernoff, J.

submitted to the EMBL Data Library, April 1995

A;Description: Human p21-activated protein kinases regulate actin organization in mammal

A;Reference number: S58682

A;Accession: S58682

A;Molecule type: DNA

A;Residues: 1-525 <SEL>

A;Cross-references: UNIPROT:Q13177; EMBL:U24153; NID:g780807; PIDN:AAA65442.1; PID:g7808

R;Martin, G.A.; Bollag, G.; McCormick, F.; Abo, A.

EMBO J. 14, 1970-1978, 1995

A;Title: A novel serine kinase activated by rac1/CDC42Hs-dependent autophosphorylation i

A;Reference number: S55258; MUID:95262637; PMID:7744004

A;Accession: S55258

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 'MEETQCKSNLEL', 31-89, 'T', 91-149, 'F', 151-224, 'T', 226-328, 'R', 330-338, 340-525

A;Accession: S55304

A;Molecule type: protein

A;Residues: 402-418 <MAW>

R;Martin, G.A.; Bollag, G.; McCormick, F.; Abo, A.

EMBO J. 14, 4385, 1995

A;Reference number: S58690; MUID:96016211; PMID:7556080

A;Contents: erratum

A;Accession: S58690

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-30 <MAP>

R;Benner, G.E.; Dennis, P.B.; Masaracchia, R.A.

J. Biol. Chem. 270, 21121-21128, 1995

A;Title: Activation of an S6/H4 kinase (PAK 65) from human placenta by intramolecular an

A;Reference number: A57441; MUID:95403344; PMID:7673144

A;Accession: A57441

A;Molecule type: protein

A;Residues: 197-216;402,'S',404-409 <BEN>
A;Experimental source: placenta
C;Superfamily: serine/threonine protein kinase
C;Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/threonine
F;244-501/Domain: protein kinase homology <KIN>
F;255-263/Region: protein kinase ATP-binding motif
F;197,402/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 6.0%; Score 305.5; DB 2; Length 525;
Best Local Similarity 28.1%; Pred. No. 4.9e-07;
Matches 94; Conservative 54; Mismatches 160; Indels 27; Gaps 8;

Qy 320 PSLSRGAHEKPSVEYLHVALQGSVSSQASHLSLAKTAAAGSRGSRSPKTEDNEG 379
Db 181 PVIAIRPDHTKSIYTRSDVPAPVFGSHVDG-----AAKSLDKQKKPKMTDEE- 231
Qy 380 VLLTEKLKPVDYVEEVEVHMATHQLRGSGFGEVHRMEDKQTQPCQAVKVRLEVFRAE 439
Db 232 --IMEKLRTI-VSIGDPKKKYTRYEKIQGASGTVFTATDVALGQEVAKQINLQKQPKK 288
Qy 440 EL-----MACAGLTSPIRVLYGAVREGPWNIWFELLEGSLGOLVKEQGLPEDRALY 494
Db 289 ELIINEILLMKLPNINVFLDQSYLVGDELPVVMVEYLAGSLTDVVTETACMDAQIAA 348
Qy 495 YLGOALEGLEYLHSRIILHGDVKADNVLLSSDGSAAALCDFGHAVCLQPDGGLGKSLTGD 554
Db 349 VCRCLOALEFLHANQVHROIKSDNVLLGMEGS-VKLTDFGCAQIITPEQSKRSTMV-- 405
Qy 555 YIPGTHMAPBEVLGRSCDAKVQVWSSCCMMLHMLNGCHPWQTFRRPL--CLKIASGP 612
Db 406 ---GTFYMAPEVVRKAYGPKVYDWSLGIAMENVEGEPYVYN--ENPLRALYLIATNG 460
Qy 613 PPVRIIPSCAPLTAQATQEGLEKPEPIHRVSAEL 647
Db 461 TPELQNPKEKLSIFRDLNRCLEMDVEKRGSAKEL 495

RESULT 14
T36746
probable serine/threonine protein kinase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36746
R;Saunders, D.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A;Reference number: Z21613
A;Accession: T36746
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-550 <SAU>
A;Cross-references: UNIPROT:Q9S2A6; EMBL:AL096849; PIDN:CAB50939.1; GSPDB:GN00070; SCOE
A;Experimental source: strain A3(2)
C;Superfamily: serine/threonine protein kinase, PksC type; protein kinase homology

Query Match 6.0%; Score 305; DB 2; Length 550;
Best Local Similarity 31.7%; Pred. No. 5.4e-07;
Matches 106; Conservative 37; Mismatches 135; Indels 56; Gaps 13;

Qy 406 LGRGSFGEVHRMEDKQTQPCQAVKVRV-----EVFR-----AEELMACAGLTSPIRV 453
Db 21 LLEGGMGTVWRARDEVLRREVAVKEVRAPAGLSQPDVGRMYARLEREAARISHPNVV 80
Qy 454 PLYGAVREG--PWNIWFELLEGSLGOLVKEQGLPEDRALYILGOALEGLEYLHSRI 511
Db 81 TYVDVATDGGREW--IVMELVRGLSLADLLDAEGPLPRRAALIGAEVLAAALRAAHAAGV 138
Qy 512 LHGDVKADNVLLSSDGSAAALCDFGHAVCLQPDGGLGKSLTGDYIPGTHMAPBEVLGR 571
Db 139 LHRDVKPANVLLANDG-RVLLTDFGIA---RVESGSEALTMTEGVV-GSPEFLAPRALGR 193
Qy 572 SCDAKVDVWSSCCMMLHMLNGCHPWQTFRRGLCLKIASPPPPVREIPPCAPLTAQATQ 631

Db 194 TPGAASDLWSLGLVLYATVEGVSPPROGTPLSTLRAIVDEAVP-----PPRRAGALGPVVE 249
Qy 632 EGLKEPEPIHRVSAELGKGNRVALQOQVG-GLKSPWGEYKEPRHPPPNQANVHOTLHAQP 690
Db 250 GLLRKDPFAERLPABE-----ABRALRLVGAGAPPGRG-----PRTGAPSGGAFAPTVVA-- 299
Qy 691 RELSPRAGPRPABETTGRAPKLOPPLPPEPPEP 724
Db 300 ----AHPG-----PPTAPTTPMP 313

RESULT 15
G96575
probable MEK kinase MAP3Ka, 84794-81452 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: G96575
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.D.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Rowley, D.; Sakano, H.
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G96575
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-608 <STO>
A;Cross-references: UNIPROT:Q9LPH2; GB:AE005173; NID:g10645340; PIDN:AAG21460.1; GSPDB:
C;Genetics:
A;Gene: F22G10.18
A;Map position: 1

Query Match 6.0%; Score 304; DB 2; Length 608;
Best Local Similarity 26.4%; Pred. No. 6.6e-07;
Matches 124; Conservative 67; Mismatches 194; Indels 84; Gaps 21;

Qy 346 SSSQAHSL-----TSLAKTAAARGSR---SREPSPKTEDNEGVLTEKLKPDYDYEYREHV 398
Db 175 SSSECHPLRPPTSPTSFSAVHSGRIGGYETSP-----SGFSTWKKG----- 218
Qy 399 WATHQLRLGRSGFGEVHRMEDKQTQFCQAVKVRV-----EVFRA--BELMACAGLIT 448
Db 219 -----FLGSGTFGQVILGFNSKGMCAIKEVKVISDDQTSKECLKQLNQEINLNLQIC 272
Qy 449 SPRIVLYGAVREGPWNIWFELLEGSLGOLVKEQGLPEDRALYILGOALEGLEYLHS 508
Db 273 HPIVQYVGSSELSEETLSVILEYVSGSIHKLKDYGSFTFPIQNYTROIAGLAYLHG 332
Qy 509 RRIILHGDVKADNVLLSSDGSAAALCDFG---HAVCLQPDGGLGKSLTGDYIPGTHMAP 565
Db 333 RNTVHRDIKGANILVDPNG-EIKLADFQWAKHVTAF-----STMLS---FKGSPYVMAP 382
Qy 566 EVLGRS-CDAKVDVWSSCCMMLHMLNGCHPWQTFRRGLCLKIA-SPPPPVREIPPCSA 623
Db 383 EVWVSQNGYTHAVDIWSLGTCTILENATSKPPWQFEGVAAIFKIGNSKDTP--EIPDHL 440
Qy 624 PLTAQAIQEGLEKPEPIHRVSAE-LGGKVNALQOVGLKSP-----WRGEYKEP-RH 674
Db 441 NDAKNFILCLQNRNTPVPTASQLLEHFLNTRVASTSLPKDPFRPSYDGNFSLPTRE 500
Qy 675 PPPNQANYHOTLHAQPRELSPAPGR-----PAEETTGRAPKLOPPL-----PPE 720
Db 501 PYPGRLS-HDNVAKQPLSRTIKSPRENVRATISLPVSPCSPLQLGPAYKSCFLSPH 559
Qy 721 PPE--PNKSPPLTUSKEESGWPEPLSLSEAPAPAR--NPSSPERKATV 765
Db 560 PSYAPFGQDSGYNLAEFAASPPRMKKDAMMEFSSFTQTPTFNSPLSRV 608

Search completed: June 7, 2005, 12:03:33
Job time : 37.0549 secs

THIS PAGE BLANK (USPTO)